CANADIAN JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. II - NO. 4

1950

Published for

THE CANADIAN MATHEMATICAL CONGRESS

by the

University of Toronto Press

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, H. S. M. Coxeter, University of Toronto. Every paper should contain an introduction summarizing the results as far as possible in such a way as to be understood by the non-expert.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is \$6.00. This is reduced to \$3.00 for individuals who are members of the following Societies:

Canadian Mathematical Congress
American Mathematical Society
Mathematical Association of America
London Mathematical Society
Société Mathématique de France

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of British Columbia
Carleton College
Ecole Polytechnique
Loyola College
University of Manitoba
McGill University
McMaster University
Universite de Montreal
Queen's University
Royal Military College
University of Toronto
National Research Council
and the
American Mathematical Society

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

CONTRIBUTION A L'ETUDE DU PROBLEME DES TIMBRES POSTE

JACQUES TOUCHARD

On demande le nombre S_n de manières dont on peut replier une bande de n timbres-poste (TP) sur un seul timbre. Dans le graphique de Sainte Lagüe (graphique SL) [3, p. 39], les liaisons entre timbres ne doivent pas se couper (fig. 1). Nous supposerons toujours que le timbre 1 occupe la première place à gauche, car si P_n est, dans ce cas, le nombre de permutations bonnes, on aura $S_n = nP_n$. Nous supposerons de plus que la liaison (1, 2) est au-dessus de l'axe. Nous appellerons arc pair ou couple pair un arc $(\alpha, \alpha + 1)$ où α est pair; arc impair ou couple impair un arc $(\alpha, \alpha + 1)$ où α est impair. Nous appellerons aussi permutation inverse de $g = |1, a, b, \ldots, p, q|$, la permutation $g' = |1, q, p, \ldots, b, a|$.

Dans le §1 de ce travail, j'indique une méthode qui divise le problème en plusieurs autres. Les §§2 et 3 m'ont été inspirés par la lecture d'une très intéressante brochure de M. Albert Sade, parue récemment. Le §3 met en jeu, pour la première fois dans cette question, des groupes d'operations. Dans le §4, j'aborde, sans parvenir à le résoudre, un problème de configurations plus simple que celui des TP. La méthode suivie contient une notion, celle des systèmes propres, qui pourra, je pense, être utilisée dans le problème des TP lui-même et je crois aussi que la solution du problème plus simple, en dehors de son intérêt propre, pourrait ouvrir une voie toute différente et moins épineuse pour traiter le problème des TP. Le §5 donne quelques tables numeriqués. Les valeurs de P_n, jusqu'à $n = 10$, ont été obtenues par M. H. W. Becker, d'Omaha, Nebraska et par moi-même. Celles de P_{11} et de P_{12} sont dûes exclusivement à M. A. Sade.

J'ai rassemblé dans ce mémoire l'essentiel de ce que je connais sur le problème des TP, en laissant toutefois de côté la représentation de certaines configurations par des substitutions de n lettres, qui exigerait de longs développements.

Sauf dans les §§2 et 3, je n'ai presque pas donné de démonstrations, celles-ci étant fort longues et exigeant de nombreuses figures, et aussi parce que j'espère pouvoir revenir sur la question avec des résultats plus avancés.

1. En nous référant au graphique SL, nous appellerons point double de première espèce l'intersection de deux arcs, $(\alpha, \alpha + 1) \times (\beta, \beta + 1)$, de même parité, $\beta - \alpha = 2, 4, \ldots.$ Rabattons le demi-plan inférieur à l'axe sur le

Reçu le 7 juin, 1949.

demi-plan supérieur, nous obtenons un graphique (Γ), fig. 2. Nous appellerons point double de 2^e espèce ou point double fictif (p.d.f.) l'intersection de deux arcs $(\alpha, \alpha + 1) \times (\beta, \beta + 1)$ de parité différente, $\beta - \alpha = 3, 5, 7, \ldots$. Le problème des TP revient à chercher quel est le nombre des permutations qui ne contiennent aucun point double de première espèce, que nous appellerons permutations bonnes, et c'est, en somme, un des problèmes fondamentaux que pose la théorie des permutations. Dans le graphique (Γ) nous supposerons toujours que la permutation est bonne; on n'aura donc pas de point

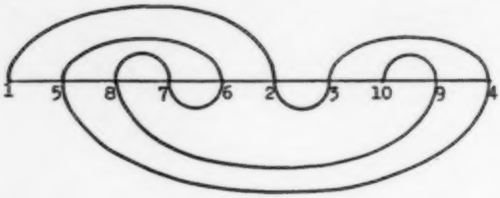

FIGURE 1

double de 1^e espèce, mais on peut avoir $0, 1, 2, \ldots$ p.d.f. et, si T_k^n est le nombre des figures qui présentent k p.d.f., on aura

$$P_n = T_0^n + T_1^n + T_2^n + \ldots .$$

On a évidemment

$$T_0^n = 2^{n-2} \tag{1}$$

puisque dans le graphique (Γ), il y a toujours deux places disponibles pour un nouveau timbre. Lorsqu'il y a k p.d.f., il se peut qu'aucun des arcs $(1, 2) - (2, 3) - (3, 4) - \ldots - (\alpha-1, \alpha)$ ne soit coupé et que l'arc $(\alpha, \alpha+1)$ soit coupé. Dans ce cas, il est coupé par un ou plusieurs arcs $(\beta, \beta + 1)$, $(\gamma, \gamma + 1), \ldots$ où $\beta > \alpha + 1$, $\gamma > \alpha + 1, \ldots$. On peut alors assimiler l'ensemble des timbres $\alpha + 1$, $\alpha + 2$, $\alpha + 3$, $\ldots$ à un seul $(\alpha + 1)^{\text{ème}}$ timbre et, d'après (1), il y a $2^{\alpha+1-2} = 2^{\alpha-1}$ dispositions sans p.d.f. des timbres $1, 2, \ldots, \alpha + 1$. Supprimons maintenant les timbres $1, 2, \ldots, \alpha - 1$; le timbre α devient le premier timbre et nous avons la proposition suivante: si a_k^n est le nombre des figures du graphique (Γ), relatives à n timbres, dans lesquelles le premier arc est coupé, on a

$$T_k^n = a_k^n + 2\, a_k^{n-1} + \ldots + 2^h\, a_k^{n-h} + \ldots . \tag{2}$$

On peut raisonner de même sur les deux timbres extrêmes $n - 1$ et n, ce qui donnera pour a_k^n une expression entièrement analogue à (2) et l'on démontre ainsi le théorème suivant:

$$T_k^n = d_k^n + 2.2\, d_k^{n-1} + \ldots + i\,.\,2^{i-1}\, d_k^{n-i+1} + \ldots ,$$

où d_k^n est le nombre des figures du graphique (Γ), relatives à n timbres et

dans lesquelles le premier et le dernier arcs sont coupés. Ce théorème a une signification physique si l'on se représente une pile de TP repliés sur le premier. Il simplifie un peu la recherche de T_k^n et l'on trouve

$$T_1^n = (n-4)\,2^{n-4} + (n-6)\,2^{n-6} + (n-8)\,2^{n-8} + \dots,$$

$$T_2^n = \frac{n-1}{2}\,[(n-6)\,2^{n-6} + 2(n-8)\,2^{n-8} + \dots + (i-2)(n-2i)\,2^{n-2i} + \dots].$$

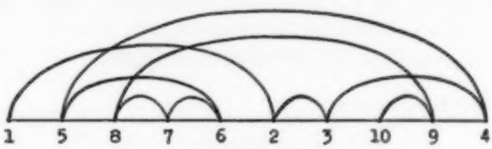

FIGURE 2

On a aussi

$$9\,T_1^n = (3n-14)\,2^{n-3} + 9 + (-1)^{n-1},$$

$$108\,T_2^n = (n-1)[(3n-22)\,2^{n-1} + 27(n-1) + (-1)^{n-1}(3n-11)].$$

La recherche de T_3^n serait plus difficile et déjà celle de T_2^n exige des précautions, parce qu'il arrive que l'existence de 2 p.d.f. entraîne obligatoirement celle d'un troisième et même d'un quatrième p.d.f.

Je n'ai pas recherché d'une manière définitive le nombre maximum de p.d.f. que peut présenter une permutation bonne. Je note seulement ceci:

Soit $n = 4\lambda + 1$, la permutation

$$|1|4, 8, \dots, n-1|n-2, n-6, \dots, 3|2, 6, \dots, n-3|n, n-4, \dots, 9, 5|.$$

Soit $n = 4\lambda + 3$, la permutation

$$|1|4, 8, \dots, n-3|n, n-4, \dots, 3|2, 6, \dots, n-1|n-2, n-6, \dots, 9, 5|$$

et leurs inverses donnent chacune $\frac{1}{8}(n-1)(n-3)$ p.d.f. J'ai divisé ces permutations en tranches qui sont des progressions arithmétiques de raison $+4$ ou -4.

Soit $n = 4\lambda$, la permutation

$$|1|4, 8, \dots, n-8|n-1, n-4, n-5, n|n-9, n-13, \dots, 3|$$
$$|2, 6, \dots, n-6|n-3, n-2|n-7, n-11, \dots, 9, 5|.$$

Soit $n = 4\lambda + 2$, la permutation

$$|1|4, 8, \dots, n-6|n-3, n-2|n-7, n-11, \dots, 3|2, 6, \dots, n-8|$$
$$|n-1, n-4, n-5, n|n-9, n-13, \dots, 9, 5|$$

et leurs inverses donnent chacune $\frac{1}{8}(n-2)(n-4) + 1$ p.d.f. J'ai divisé ces

permutations en 7 tranches; la 2^e, la 4^e, la 5^e et 7^e sont des progressions arithmétiques de raison $+4$ ou -4. D'après cela, le mathématicien qui résoudra le problème des TP doit vraisemblablement s'attendre à voir apparaître des séries dans le genre de $\sum_0^\infty a_n z^n x^{n^2}$, c'est-à-dire des séries qui se présentent dans la théorie des fonctions elliptiques ou dans la théorie des partitions.

2. En revenant au graphique SL, M. A. Sade [2] a introduit une fonction que nous appellerons $A(n, x)$, égale au nombre des permutations bonnes de n timbres, commençant par 1 et où x occupe la seconde place et il a remarqué que

$$A(n, 3) = A(n, 4) = P_{n-2}.$$

Je me propose de montrer que

$$A(n, 2k) = A(n, 2k - 1), \qquad n \geqslant 2k. \tag{3}$$

1°) *Opération* $\Omega_1(\alpha)$.

Ayant une permutation bonne $g = |1, \dots . \lambda, \mu, \alpha, \beta, \gamma, \dots|$ je laisse 1 immobile et, sur les éléments restants, je fais une substitution circulaire, de gauche à droite, à partir d'une origine α que je place immédiatement à droite de 1, de facon à obtenir $|1, \alpha, \beta, \gamma, \dots, \lambda, \mu|$. Les liaisons entre ces éléments ne sont pas modifiées et la seule chose qui puisse arriver c'est que l'arc $(1, 2)$ soit coupé par un ou plusieurs arcs impairs. Or ceci n'arrivera pas si, entre 1 et α, origine de la substitution circulaire, il y a zéro ou un nombre entier de couples impairs. Cette condition est nécessaire et suffisante. Si n est impair, la position du dernier timbre n est indifférente, car, alors, l'arc impair $(n, n+1)$ n'existe pas.

2°) *Opération* $\Omega_2(\alpha)$.

Même définition que pour $\Omega_1(\alpha)$, la substitution circulaire étant faite de droite à gauche, de facon à obtenir: $|1, \alpha, \mu, \lambda, \dots, \gamma, \beta|$. Appelons conjugué α' de α le nombre $\alpha' = 2k$, si $\alpha = 2k - 1$ et $\alpha' = 2k - 1$, si $\alpha = 2k$. Pour que la permutation $g\Omega_2(\alpha)$ soit bonne, il faut et il suffit qu'à gauche de α se trouvent son conjugué α' et zéro ou un nombre entier de couples impairs. Si n est impair, la place du timbre n est indifférente. L'opération Ω_2 revient à faire l'opération Ω_1 sur la permutation inverse de g. Comme exemple, soit la permutation bonne

$$g = |1, 6, 5, 4, 3, 2, 7, 8, 11, 10, 9, 12|.$$

On a:

$$\begin{aligned}
g\Omega_1(6) &= g,\\
g\Omega_1(4) &= |1, 4, 3, 2, 7, 8, 11, 10, 9, 12, 6, 5|,\\
g\Omega_1(2) &= |1, 2, 7, 8, 11, 10, 9, 12, 6, 5, 4, 3|,\\
g\Omega_1(7) &= |1, 7, 8, 11, 10, 9, 12, 6, 5, 4, 3, 2|,\\
g\Omega_1(11) &= |1, 11, 10, 9, 12, 6, 5, 4, 3, 2, 7, 8|,
\end{aligned}$$

qui sont toutes bonnes, de même que $g\Omega_2(5)$, $g\Omega_2(3)$, $g\Omega_2(2)$, $g\Omega_2(8)$ et $g\Omega_2(12) = g'$, inverse de g.

Pour l'opération $\Omega_1(a)$, on peut toujours prendre comme origine a le timbre 2 et aussi le timbre situé immédiatement à droite de 2. Pour l'opération $\Omega_2(a)$, on peut toujours prendre comme origine a le timbre 2, le timbre situé immédiatement à gauche de 2 et le dernier timbre à droite. Il y a P_{n-1} permutations bonnes où 2 occupe la place 2 et P_{n-1} permutations bonnes où 2 occupe la place n. On a donc la proposition suivante:

Parmi les $P_n - 2P_{n-1}$ permutations bonnes ou 2 n'occcupe ni la place 2 ni la place n, il y a au moins 3 permutations qui présentent le même ordre circulaire de leurs éléments, et au moins 3 permutations qui présentent l'ordre circulaire inverse.

Quant à l'égalité (3), sa démonstration est immédiate. En effectuant l'opération $\Omega_2(2k)$ sur les permutations $A(n, 2k - 1)$, on obtient les $A(n, 2k)$ et, en effectuant l'opération $\Omega_2(2k - 1)$ sur les permutations $A(n, 2k)$ on obtient les $A(n, 2k - 1)$. Il est clair, en effet, que si $2k - 1$ occupe la 2^e place, $2k - 1$ et $2k$ se trouvent sous l'arc (1, 2) et l'arc $(2k - 1, 2k)$ recouvre zéro ou un nombre entier d'arcs impairs. Il y a donc correspondance "one-one" entre les permutations $A(n, 2k)$ et les permutations $A(n, 2k - 1)$.

3. M. Sade a également introduit une fonction que nous appellerons $B(n, i)$ et qui, dans le graphique SL, relatif à n timbres est égale au nombre i de places disponibles pour un $(n + 1)^{ème}$ timbre. Il a montré que le maximum r de i est $r = \nu + 1$, si $n = 2\nu$ ou $2\nu + 1$; que $B(2\nu, r) = 2^{\nu-1}$, $B(2\nu + 1, r) = 2^{\nu}$; et il a imaginé un procédé pour former les permutations bonnes, donnant le maximum r de places disponibles. Nous modifierons et complèterons ce procédé et adopterons les définitions suivantes:

Soit g une permutation des nombres $1, 2, 3, \ldots, n$, commencant par 1. Si, dans g, il existe un ensemble E formé par $\beta - \alpha + 1$ nombres successifs $\alpha, \alpha + 1, \alpha + 2, \ldots, \beta$, dans un ordre quelconque, nous désignerons par $\rho(\alpha, \beta)$ l'opération qui consiste à renverser l'ordre des éléments de E, sans modifier ceux qui précèdent ou qui suivent E. Par exemple, si

$$g = |1, 2, 3, 9, 10, 8, 7, 5, 6, 4|,$$

on aura

$$\begin{aligned} g\rho(7, 10) &= |1, 2, 3, 7, 8, 10, 9, 5, 6, 4|, \\ g\rho(4, \ \ 7) &= |1, 2, 3, 9, 10, 8, 4, 6, 5, 7|. \end{aligned}$$

L'opération $\rho(\alpha, \beta)$ ne serait pas définie si, entre des éléments de l'ensemble E, se trouvaient des nombres $< \alpha$ ou $> \beta$. Lorsque $\beta = n$, j'écrirai plus simplement $\rho(\alpha, n) = \rho(\alpha)$ et l'opération $\rho(\alpha)$ consiste à renverser l'ordre de tous les nombres $\geqslant \alpha$, supposés réunis dans un même ensemble. L'opération $\rho(1)$ est exclue.

Cela étant, partant de la permutation naturelle $H = |1, 2, 3, \ldots, n|$, les opérations $\rho(2)$, $\rho(3)$, $\rho(4), \ldots, \rho(n - 1)$ et $\rho(n) = 1$ forment la base d'un

groupe abélien G_0, d'ordre 2^{n-2}, dont toutes les opérations sont d'ordre 2. L'opération $\rho(\alpha)$, appliquée à H, revient à transposer α et l'ensemble $\alpha + 1$, $\alpha + 2, \ldots, n$ en renversant l'ordre des éléments de cet ensemble. G_0 est donc isomorphe à un groupe de substitution de $2n - 4$ lettres dont les éléments générateurs sont $n - 2$ transpositions sans lettres communes.

L'ensemble des permutations HG_0 est formé par les permutations sans point double fictif du §1, car aucune opération de G_0 ne peut créer de p.d.f. si elle est appliquée à une permutation sans p.d.f.

En prenant un certain nombre d'operations $\rho(a_1)$, $\rho(a_2), \ldots, \rho(a_q)$ et l'opération identique, on forme un sous-groupe invariant de G_0 d'ordre 2^q.

En ce qui concerne le nombre i de places disponibles pour le $(n + 1)^{ème}$ timbre, la permutation H en donne évidemment le nombre maximum $i = r$. Or on verra facilement que:

(A) l'opération $\rho(2)$ n'enlève aucune place disponible.

(B) si $n = 2\nu$, les opérations $\rho(2p)$ n'en enlèvent aucune; l'opération $\rho(2p + 1)$ enlève p places disponibles et si, dans une opération du groupe G_0, figure le produit

$$\rho(2k_1 + 1)\,\rho(2k_2 + 1) \ldots \rho(2k_q + 1), \qquad k_1 < k_2 < \ldots < k_q,\ k_q \neq 0$$

il y aura, pour le $(n + 1)^{ème}$ timbre, $i = r - k_q = \nu + 1 - k_q$ places disponibles.

(C) si $n = 2\nu + 1$, les opérations $\rho(2p + 1)$ n'enlèvent aucune place disponible; l'opération $\rho(2p)$ enlève $p - 1$ places et si, dans une opération de G_0, figure le produit

$$\rho(2k_1)\ \rho(2k_2) \ldots \rho(2k_q), \qquad k_1 < k_2 < \ldots < k_q,\ k_q \neq 0$$

il y aura, pour le $(n + 1)^{ème}$ timbre, $i = r - k_q + 1 = \nu + 2 - k_q$ places disponibles.

De sorte que, parmi les 2^{n-2} permutations sans p.d.f. du §1, il y en a:

si $n = 2\nu$,	$2^{\nu-1}$	qui donnent $i = r = \nu + 1$	
et	$2^{\nu+m-1}$	qui donnent $i = \nu - m$	$(m = 0, 1, 2, \ldots, \nu - 2)$,
si $n = 2\nu + 1$,	2^{ν}	qui donnent $i = r = \nu + 1$	
et	$2^{\nu+m}$	qui donnent $i = \nu - m$	$(m = 0, 1, 2, \ldots, \nu - 2)$.

Les permutations de M. Sade, qui donnent $i = r$, sont donc celles qu'on obtient en appliquant à la permutation naturelle H les opérations du sous-groupe de G_0 engendré par $\rho(2)$, $\rho(4)$, $\rho(6), \ldots, \rho(2\nu - 2)$, $\rho(2\nu)$, si $n = 2\nu$ et par $\rho(2)$, $\rho(3)$, $\rho(5), \ldots, \rho(2\nu - 1)$, $\rho(2\nu + 1)$, si $n = 2\nu + 1$.

On peut remarquer d'ailleurs et sans entrer dans le détail, que l'on obtient les permutations de M. Sade en considérant soit les couples impairs $\alpha_1, \alpha_2, \ldots, \alpha_\nu$ où $\alpha_i = (2i - 1, 2i)$ soit les couples pairs $\beta_1, \beta_2, \ldots, \beta_\nu$, où $\beta_i = (2i, 2i + 1)$ auxquels on adjoint le couple fictif $\beta_0 = (0, 1)$, et en pratiquant les opérations

d'un groupe G'_0, analogue à G_0, sur les indices des couples. En outre, il y a certaines transpositions à faire au sein des couples. C'est là la vraie raison pour laquelle le nombre i de places disponibles garde sa valeur maximum $i = r$. Les opérations du groupe G'_0 sont même les seules opérations qu'on puisse effectuer sur les couples pairs ou impairs sans créer de points doubles de première espèce.

Je serai beaucoup plus succinct en ce qui concerne les permutations qui, dans le graphique (Γ), ont un p.d.f. Il faut d'abord observer qu'appliquée à la permutation naturelle H, l'opération $\rho(\alpha, \beta)$ interdit d'effectuer les opérations $\rho(\alpha+1), \rho(\alpha+2), \ldots, \rho(\beta)$, à moins, bien-entendu, que $\beta = n$, nombre de timbres, auquel cas $\rho(\alpha, n) = \rho(\alpha)$. On a alors la proposition suivante:

Les opérations engendrées par la base

$$\rho(2), \rho(3), \ldots, \rho(\alpha), \rho(\beta+1), \rho(\beta+2), \ldots, \rho(n)$$

et $\rho(\alpha, \beta)$, où $\beta - \alpha$ est pair, forment un groupe abélien G, d'ordre $2^{n-1-\beta+\alpha}$, dont tous les éléments sont d'ordre 2. Ces opérations, appliquées à la permutation H, conservent la séquence $\alpha, \alpha+1, \alpha+2, \ldots, \beta$, dans l'ordre naturel ou dans l'ordre inverse. Celles qui ne contiennent pas $\rho(\alpha, \beta)$ ne produisent aucun p.d.f. et forment un sous-groupe de G. Celles qui contiennent $\rho(\alpha, \beta)$ produisent l'unique point double de 2^e espèce $(\alpha-1, \alpha) \times (\beta, \beta+1)$. On retrouve ainsi la valeur de $T_1{}^n$, donnée au §1.

Quant aux permutations qui offrent $r-1$ places disponibles pour le $(n+1)^{\text{ème}}$ timbre, on peut les obtenir au moyen du produit de deux opérations $\rho(\alpha, \beta)$, compatibles entre elles, et des opérations $\rho(\lambda)$, compatibles avec les précédentes. Ceci exigerait d'assez longues explications et je me bornerai à donner la valeur de $B(n, r-1)$, savoir

$$\begin{aligned} B(2\nu, r-1) &= B(2\nu, \nu) = (\nu-1)2^{\nu-1} + (\nu-2)2^{\nu-2} + (\nu-3)2^{\nu-3} + \ldots + 2, \\ B(2\nu+1, r-1) &= B(2\nu+1, \nu) \\ &= (2\nu-1)2^{\nu-1} + (2\nu-3)2^{\nu-2} + (2\nu-5)2^{\nu-3} + \ldots + 5.2^2 + 3.2. \end{aligned}$$

4. Ayant $2n$ places ou $2n$ nombres $1, 2, \ldots, 2n$, dans l'ordre naturel sur un axe, on relie les places, deux à deux, par des arcs convexes, chaque place n'étant touchée que par un seul arc. On obtient ainsi un graphique (fig. 3) analogue au graphique (Γ) du §1.

Le nombre des configurations possibles est

$$p_{2n} = 1.3.5 \ldots (2n-1)$$

et l'on demande le nombre $U_{2n}(p)$ de celles qui ont p points doubles.

On peut d'abord ranger les configurations suivant une suite bien ordonnée. On verra, en effet, que, pour $2n$ places, chaque configuration peut être représentée d'une seule manière par l'expression

$$a_2 p_{2n-2} + a_4 p_{2n-4} + \ldots + a_{2n-2} p_2 + 1, \qquad a_{2k} \leqslant 2n - 2k \tag{4}$$

qui prend toutes les valeurs entières de 1 à p_{2n}, inclusivement, grâce à l'identité

$$p_{2n} = (2n - 2)p_{2n-2} + (2n - 4)p_{2n-4} + \ldots + 2p_2 + 1.$$

L'expression (4) doit toujours se terminer à droite par l'unité, de sorte qu'il n'y a aucune ambiguité. La configuration de la fig. 3 a pour numéro le nombre $4p_8 + p_6 + p_4 + p_2 + 1 = 440$. Inversement, ayant un nombre m, $1 \leqslant m \leqslant p_{2n}$ et sachant qu'il s'agit d'une configuration de $2n$ places ou de n arcs, on mettra m sous la forme (4) et on en déduira une configuration unique.

La première idée qui vient à l'esprit serait de déterminer le nombre des points doubles d'après la valeur des coefficients a_{2p}. J'ai du y renoncer.

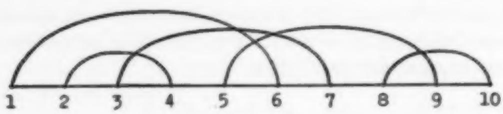

FIGURE 3

Une fois la fonction $U_{2n}(p)$ déterminée, on aura une vérification numérique, puisqu'on doit avoir l'identité

$$p_{2n} = U_{2n}(0) + U_{2n}(1) + \ldots + U_{2n}\binom{n}{2}$$

où $\binom{n}{2} = \frac{1}{2}n(n - 1)$ est le nombre maximum de points doubles. C'est là une vérification qui ne se présente pas dans le problème des TP, puisque P_n n'est pas connu, sauf pour les premières valeurs de n.

Posons

$$(5) \qquad f_p(x) = U_0(p) + U_2(p)x^2 + U_4(p)x^4 + \ldots + U_{2n}(p)x^{2n} + \ldots .$$

Ce sont les fonctions $f_p(x)$ qu'il faut trouver. Il est bien connu qu'en faisant par convention $U_0(0) = 1$, on a

$$2x^2 f_0(x) = 1 - (1 - 4x^2)^{\frac{1}{2}}$$

et les nombres $U_{2n}(0)$ sont appelés nombres de Segner. Pour rechercher $f_p(x)$, $p > 0$, j'aurai recours à la notion de système d'arcs que M. A. Errera [1] a imaginée dans le cas des configurations sans points doubles. En généralisant l'idée de cet auteur, nous dirons que deux arcs C_1 et C_2 appartiennent à un même système si l'un recouvre l'autre ou si l'un coupe l'autre, ou si un troisième arc C_3 recouvre C_1 et C_2 ou les coupe tous les deux ou, encore, coupe l'un d'eux et recouvre l'autre. On peut dire aussi qu'un système est un ensemble d'arcs, reliant des points deux à deux et tel que tout point du système, sauf ses deux points extrèmes, soit recouvert par un arc.

Lorsqu'un système S_1 est recouvert par un arc d'un système S_2 et qu'aucun arc de S_1 n'est coupé par aucun arc de S_2, S_1 et S_2 forment un système S, dont S_1 est un sous-système.

Nous dirons qu'un système est *propre*, lorsqu'il ne contient pas de sous-

système. Un système à p points doubles est propre à $2n$ places, lorsqu'il figure dans les configurations de $2n$ places et qu'il ne figure pas dans les configurations de $2n-2$ places. Soit

$$\sigma_{2n}(p) \tag{6}$$

le nombre des systèmes propres à p points doubles et $2n$ places, on a $\sigma_{2n}(p)=0$ si $2n > 2p+2$. En effet le système

$$(1,3)(2,5)(4,7)(5,8)\ldots(2n-4,2n-1)(2n-2,2n)$$

présente $n-1$ points doubles et l'on reconnait aisément qu'il ne peut exister de système propre plus long que celui-là. On peut appeler systèmes longs les systèmes propres à p points doubles et à $2p+2$ places. Comme, d'autre part, le nombre maximum de points doubles est $\frac{1}{2}n(n-1)$, il n'existe de systèmes propres à p points doubles que pour

$$1+(1+8p)^{\frac{1}{2}} \leqslant 2n \leqslant 2p+2.$$

J'aurai maintenant besoin des séries suivantes:

$$g_p(y) = \sigma_2(p)y^2 + \sigma_4(p+1)y^4 + \ldots + \sigma_{2\mu}(p+\mu-1)y^{2\mu} + \ldots, \tag{7}$$

$$G(y,z) = g_0(y) + z^2 g_1(y) + \ldots + z^{2k} g_k(y) + \ldots, \tag{8}$$

où y et z sont deux variables indépendantes quelconques et nous poserons en outre

$$y(x,z) = xf_0(x)z + \ldots + xf_n(x)\,z^{2n+1} + \ldots. \tag{9}$$

Cela étant, on trouve, par un calcul assez long, mais sans difficulté, que

$$\begin{aligned}
f_1 &= \sigma_2(0)x^2(2f_0f_1) + \sigma_4(1)x^4f_0^4,\\
f_2 &= \sigma_2(0)x^2(2f_0f_2 + f_1^2) + \sigma_4(1)x^4(4f_0^3f_1) + \sigma_6(2)x^6f_0^6,\\
f_3 &= \sigma_2(0)x^2(2f_0f_3 + 2f_1f_2) + \sigma_4(1)x^4(4f_0^3f_2 + 6f_0^2f_1^2)\\
&\qquad + \sigma_6(2)x^6(6f_0^5f_1) + \sigma_6(3)x^6f_0^6 + \sigma_8(3)x^8f_0^8,\\
f_4 &= \quad \ldots, \qquad \ldots,
\end{aligned}$$

c'est-à-dire que, si l'on substitue la série (9) dans (7) et (8),

(10) $f_p(x)$ est, pour $p = 1, 2, 3, \ldots$ le coefficient de z^{2p+2} dans $G(y, z)$.

Or, d'une part, d'après (9),

$$zy(x,z) = xf_0z^2 + x(f_1z^4 + f_2z^6 + f_3z^8 + \ldots); \tag{11}$$

d'autre part, si on développe $G[y(x,z),z]$ suivant les puissances de z^2, on aura d'abord le terme $x^2f_0^2z^2$, puis, d'après (10), les termes $f_1z^4 + f_2z^6 + f_3z^8 + \ldots$ de sorte que

$$G[y(x,z),z] = x^2f_0^2z^2 + f_1z^4 + f_2z^6 + f_3z^8 + \ldots, \tag{12}$$

et, en comparant (11) et (12) et remarquant que $x^2f_0^2 = f_0 - 1$, on obtient l'équation, d'apparence très simple,

$$(13) \qquad zy = xz^2 + xG(y, z),$$

qui définira la fonction $y(x, z)$, lorsque la fonction $G(y, z)$, où y et z sont maintenant deux variables indépendantes quelconques, sera connue. L'équation (13) est fondamentale dans cette théorie. La racine y, qui s'annule avec x, pourra, sous réserve de la question de convergence, être développée par la série de Lagrange.

Ainsi, la détermination de $U_{2n}(p)$ est ramenée à celle des séries $g_p(y)$, c'est-à-dire à celle du nombre des système propres. Cette détermination est très difficile, tout au moins pour l'auteur de ce mémoire. Les calculs sont d'une extrême complication et je ne suis pas encore parvenu à un résultat général. Une des difficultés du problème est qu'on manque, en quelque sorte, de données expérimentales qui permettraient de rectifier des erreurs presque inévitables lorsqu'on tient compte de la complexité des figures. En effet, déjà pour $2n = 12$, le nombre des configurations dépasse dix mille et il est pratiquement impossible de tracer dix mille figures.

La détermination la plus facile est celle du nombre $\sigma_{2p}(p - 1)$ des systèmes propres longs, grace à la propriété suivante: lorsque deux arcs C_1 et C_2 sont coupés par un même arc C_3, un quatrième arc C_4 ne peut couper que C_1 ou que C_2, car si C_4 les coupait tous les deux, on aurait 2 points doubles de plus et seulement 2 places de plus, de sorte que l'on n'aurait plus un système long. On trouve ainsi que $g_0(y)$ satisfait à l'équation

$$(14) \qquad g_0^2(y) - y^2 g_0(y) + y^4 = 0,$$

et, en développant la racine qui s'annule avec y par la série de Lagrange, on obtient

$$\sigma_{2p}(p - 1) = \frac{(3p - 3)!}{(p - 1)!(2p - 1)!}, \qquad p \geqslant 1.$$

La détermination de $g_1(y)$ est plus ardue. On trouve que $g_1(y)$ est une fonction rationnelle de $g_0(y)$ et on peut la mettre sous la forme, utilisée en algèbre dans la transformation de Tschirnhausen et qui est ici

$$(15) \qquad (4y^2 - 27y^4)g_1(y) = 9y^6 + (4y^2 - 30y^4)g_0(y) + (25y^2 - 4)g_0^2(y).$$

On peut obtenir ainsi une équation aux différences pour les nombres $\sigma_{2p}(p)$, mais il se trouve que si l'on pose $u(y) = g_0(y)/y^2$, l'équation (15) est satisfaite par

$$g_1(y) = \frac{y^4}{3} \frac{d}{dy} (y^3 u^6)$$

de sorte que

$$\sigma_{2p}(p) = \frac{(3p - 2)!}{(2p + 1)!(p - 3)!} - \frac{(3p - 4)!}{(2p + 1)!(p - 5)!}$$

$$= 2(2p - 3) \frac{(3p - 4)!}{(2p)!(p - 3)!}.$$

L'équation du 3^e degré à laquelle satisfait $g_1(y)$ est compliquée et je ne l'écrirai pas. A cet égard, il serait peut-être utile d'avoir recours à la fonction $\wp u$ de Weierstrass, avec les invariants $g_2 = 4y^2$, $g_3 = -4y^4$, c'est-à-dire de poser $g_0(y) = \wp(u; 4y^2, -4y^4)$ de sorte que l'équation (14) prendrait la forme très simple $\wp'(u) = 0$. C'est un artifice que je n'ai pas encore pu essayer.

Avec les valeurs que je possède actuellement des $\sigma_{2n}(p)$, je serais théoriquement en mesure de calculer l'expression des fonctions $f_p(x)$ jusqu'à $p = 6$ inclusivement. J'ai déjà donné l'expression de $f_0(x)$. On a de plus

$$2x^4f_1(x) = 2x^2 - 1 + (1 - 4x^2 + 2x^4)(1 - 4x^2)^{-\frac{1}{2}},$$

$$2x^6f_2(x) = 2x^2 - 3x^4 - 2^{-7}(320x^6 - 880x^4 + 260x^2 + 1)(1 - 4x^2)^{-\frac{1}{2}} + 2^{-7}(1 - 4x^2)^{-3/2}.$$

En développant $(1 - 4x^2)^{-\frac{1}{2}}$ et $(1 - 4x^2)^{-3/2}$, on obtient ainsi les valeurs de $U_{2n}(1)$ et de $U_{2n}(2)$ que je donnerai plus loin, mais ces calculs sont très pénibles, déjà pour $f_2(x)$. On peut heureusement s'en passer, au moins provisoirement, en se servant de l'équation fondamentale (13) et en ne conservant dans la fonction $G(y, z)$ que les termes dont on a besoin. Par exemple, pour le calcul de $U_{2n}(3)$, il suffira de réduire $G(y, z)$ à

$$\begin{aligned} P(y, z) &= \sigma_2(0)y^2 + \sigma_4(1)y^4 + \sigma_6(2)y^6 + \sigma_8(3)y^8 + \sigma_6(3)z^2y^6 \\ &= y^2 + y^4 + 3y^6 + 12y^8 + z^2y^6, \end{aligned}$$

de sorte que l'équation (13) se réduit à

$$zy = xz^2 + xP(y, z). \tag{16}$$

Si y_1 est la racine de (16) qui s'annule avec x, $f_3(x)$ sera le coefficient de z^8 dans le développement de $P(y_1, z)$ et $U_{2n}(3)$ sera le coefficient de x^{2n} dans $f_3(x)$. Comme, d'après (16),

$$P(y_1, z) = \frac{1}{x}zy_1 - z^2,$$

$U_{2n}(3)$ sera le coefficient de $x^{2n+1}z^7$ dans le développement de y_1, qu'on obtiendra, comme nous l'avons dit par la série de Lagrange.

C'est ainsi que nous avons trouvé les expressions suivantes:

$$U_{2n}(0) = \frac{1}{n}\binom{2n}{n-1},$$

$$U_{2n}(1) = \frac{1}{1}\binom{2n}{n-2},$$

$$U_{2n}(2) = \frac{1}{2}\binom{n+3}{1}\binom{2n}{n-3},$$

$$U_{2n}(3) = \frac{1}{3}\binom{n+5}{2}\binom{2n}{n-4} + \binom{2n}{n-3},$$

$$U_{2n}(4) = \frac{1}{4}\binom{n+7}{3}\binom{2n}{n-5} + \binom{n+6}{1}\binom{2n}{n-4},$$

$$U_{2n}(5) = \frac{1}{5}\binom{n+9}{4}\binom{2n}{n-6} + \left\{\binom{n+8}{2} - 1\right\}\binom{2n}{n-5} + 4\binom{2n}{n-4},$$

$$U_{2n}(6) = \frac{1}{6}\binom{n+11}{5}\binom{2n}{n-7} + \left\{\binom{n+10}{3} - \binom{n+8}{1}\right\}\binom{2n}{n-6}$$
$$+ \frac{3(3n+25)}{2}\binom{2n}{n-5} + \binom{2n}{n-4}.$$

Ces expressions sont trop peu nombreuses pour que l'on puisse essayer de deviner la forme générale de $U_{2n}(p)$. Les nombres qu'on en déduit sont en accord avec ceux de la table du §5, que nous avons obtenus en traçant les configurations jusqu'à $2n = 10$.

Pour terminer cette section, nous indiquerons des configurations particulières pour lesquelles on peut achever le calcul. Ce cas particulier s'est d'ailleurs présenté à nous dans nos tentatives pour attaquer le problème des TP par la théorie des substitutions. Appelons origine d'un arc son extrémité gauche et appelons premier arc d'un système l'arc dont l'origine est située la première à gauche. Nous dirons que tous les points doubles d'un système sont rassemblés sur le premier arc lorsque deux arcs du système ne se coupent que si l'un d'eux est le premier arc.

On demande le nombre $V_{2n}(p)$ des configurations de n arcs, à p points doubles, où, dans chaque système ou sous-système, tous les points doubles, lorsqu'il y en a, sont rassemblés sur le premier arc. On a $V_{2n}(0) = U_{2n}(0)$ et on a aussi $V_{2n}(1) = U_{2n}(1)$, car si le 1er arc d'une configuration K n'est pas coupé, il y a au moins un autre système ou sous-système K' et l'on peut raisonner sur K' comme nous venons de le faire sur K. Mais, pour $p > 1$, $V_{2n}(p)$ est différent de $U_{2n}(p)$.

Posons alors, par analogie avec (5),

$$\varphi_p(x) = V_0(p) + V_2(p)x^2 + \ldots + V_{2n}(p)x^{2n} + \ldots,$$

et, par analogie avec (9),

$$v(x, z) = x\varphi_0(x)z + x\varphi_1(x)z^3 + \ldots + x\varphi_n(x)z^{2n+1} + \ldots.$$

On constate, par un calcul sans difficulté, que $\varphi_1(x)$ est le coefficient de z^4 dans $v^2 + v^4$, c'est-à-dire dans $v^2 + v^4 + v^6 + \ldots = v^2/(1 - v^2)$, puisque ni v^6, ni $v^8, \ldots$ ne contiennent de terme en z^4; d'une façon générale, $\varphi_p(x)$ est, pour $p \geqslant 1$ le coefficient de z^{2p+2} dans

$$H(v) = \frac{v^2}{1 - v^2}.$$

Le calcul des fonctions $\varphi_p(x)$ est alors exactement le même que celui des fonctions $f_p(x)$, mais la fonction de 2 variables $G(y, z)$ est remplacée par la

fonction d'une seule variable $H(v)$. L'équation (13) est remplacée par $zv = xz^2 + xH(v)$, c'est-à-dire

$$zv^3 + x(1 - z^2)v^2 - zv + xz^2 = 0,$$

qui, en posant $v = xzw$, devient

$$w - 1 - x^2[(1 - z^2)w^2 + z^2w^3] = 0.$$

En développant la racine qui prend la valeur 1, pour $x = 0$, par la série de Lagrange, on obtient:

$$(-1)^p V_{2n}(p) = \sum_{k=0}^{k=p} \frac{(-1)^k}{n+k+1} \; \frac{(2n+k)!}{k!(n+k)!(p-k)!\,(n-p)!}.$$

Cette expression doit être nulle pour $p = n$, car il ne peut y avoir, dans ce cas, plus de $n - 1$ points doubles.

5. Tables.

VALEURS DE P_n

n	1	2	3	4	5	6	7	8	9	10	11	12
P_n	1	1	2	4	10	24	66	174	504	1406	4210	12196

VALEURS DE T_k^n

$k \backslash n$	2	3	4	5	6	7	8	9	10
0	1	2	4	8	16	32	64	128	256
1			0	2	8	26	72	186	456
2					0	6	28	112	360
3					0	2	8	54	208
4						0	2	18	80
5							0	4	28
6							0	2	14
7								0	4

VALEURS DE $U_{2n}(p)$

$2n \backslash p$	0	1	2	3	4	5	6	7	8	9	10
2	1										
4	2	1									
6	5	6	3	1							
8	14	28	28	20	10	4	1				
10	42	120	180	195	165	117	70	35	15	5	1

VALEURS DE $\sigma_{2n}(p)$

$2n$ \ p	0	1	2	3	4	5	6	7	8	9	10
2	1										
4	0	1									
6	0	0	3	1							
8	0	0	0	12	10	4	1				
10	0	0	0	0	55	77	60	35	15	5	1
12	0	0	0	0	0	273	546	570	?	?	?

Note

La rédaction du présent mémoire était achevée, lorsque divers calculs m'ont conduit à faire la remarque suivante. Parmi les configurations du §4, considérons celles où les origines des n arcs sont fixées aux points $1, 2, 3, \ldots, n$ et où leurs extrémités parcourent les $n!$ permutations des points $n+1, n+2, \ldots, 2n$. Soit, dans ce cas, $r(n, p)$ le nombre des figures qui ont p points doubles. La fonction génératrice des nombres $r(n, p)$ est, comme il est très facile de le démontrer,

$$\begin{aligned}\Pi_n(x) &= (1+x)(1+x+x^2)\ldots(1+x+x^2+\ldots+x^{n-1})\\ &= \sum r(n, p)\, x^p,\end{aligned}$$

la somme étant étendue à toutes les valeurs $p = 0, 1, 2, \ldots, \binom{n}{2}$. On a aussi

$$\Pi_n(x) = \frac{(1-x)(1-x^2)(1-x^3)\ldots(1-x^n)}{(1-x)^n}$$

et l'on voit que

$$r(n, p) = r\left(n, \binom{n}{2} - p\right).$$

Il y a là le point de départ d'une nouvelle méthode pour traiter la question du §4. Cette méthode donne la solution du problème au moyen d'une fraction continue qui se rattache aux fonctions θ de Jacobi.

RÉFÉRENCES

[1] A. Errera, *Un problème d'énumération*, Mémoires publiés par l'Académie Royale de Belgique, Collection in-8° (2) tome 11 (1931).

[2] A. Sade, *Sur les chevauchements des permutations*, Edité chez l'auteur (14 boulevard du Jardin Zoologique, Marseille, France, 1949).

[3] A. Sainte Lagüe, *Les réseaux ou graphes*, Mémorial des Sciences Mathématiques, fascicule 18 (1926).

Lausanne, Suisse

ON REPRESENTATIONS AS A SUM OF CONSECUTIVE INTEGERS

W. J. LEVEQUE

1. Introduction. It is the object of this paper to investigate the function $\gamma(m)$, the number of representations of m in the form

$$(r+1)+(r+2)+\ldots+s, \tag{1}$$

where $s > r \geqslant 0$. It is shown that $\gamma(m)$ is always equal to the number of odd divisors of m, so that for example $\gamma(2^k) = 1$, this representation being the number 2^k itself. From this relationship the average order of $\gamma(m)$ is deduced; this result is given in Theorem 2. By a method due to Kac [2], it is shown in §3 that the number of positive integers $m \leqslant n$ for which $\gamma(m)$ does not exceed a rather complicated function of n and ω, a real parameter, is asymptotically $nD(\omega)$, where $D(\omega)$ is the probability integral

$$(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\omega} e^{-\frac{1}{2}x^2}\, dx.$$

In §4, these theorems are extended to $\gamma(m, s)$, the number of representations of m as the sum of positive consecutive terms in any of the s arithmetic progressions having constant difference s.

2. The average order of $\gamma(m)$. First we prove

THEOREM 1. *$\gamma(m) = \tau(\bar{m})$ where $\tau(u)$ is the number of divisors of u and* $m = 2^{a-1}\bar{m}$, $\bar{m}$ odd.

For by (1) we have

$$m = \frac{s^2+s}{2} - \frac{r^2+r}{2}, \quad 2m = (s-r)\,(s+r+1).$$

Putting $s - r = n$, this gives

$$2m = n(n+2r+1).$$

Since n and $n+2r+1$ have opposite parity, and since $n < (2m)^{\frac{1}{2}}$, $\gamma(m)$ is the number of ways of writing $2m$ as the product of an even and an odd number. That is,

$$\gamma(m) = \sum_{\substack{n \mid \bar{m} \\ n < (2m)^{\frac{1}{2}}}} 1 + \sum_{\substack{2m/n \mid \bar{m} \\ 2m/n > (2m)^{\frac{1}{2}}}} 1 = \sum_{d \mid \bar{m}} 1 = \tau(\bar{m}).$$

THEOREM 2. *The average order of $\gamma(m)$ is $\frac{1}{2}\log m$; more precisely,*

$$\frac{1}{n}\sum_{m=1}^{n} \gamma(m) = \frac{1}{2}\log n + \frac{2C + \log 2 - 1}{2} + O\,(n^{-\frac{1}{2}}),$$

where C is Euler's constant.

Received April 1, 1949.

For let l be the unique integer such that $2^l \leqslant n < 2^{l+1}$. Then by Theorem 1,

$$\sum_{m=1}^{n} \gamma(m) = \sum_{m=1}^{n} \tau(\overline{m})$$

$$= \sum_{\substack{1 \leqslant m \leqslant n \\ m \equiv 1 \pmod 2}} \tau(m) + \sum_{\substack{1 \leqslant m \leqslant n \\ m \equiv 2 \pmod 4}} \tau(m/2) + \sum_{\substack{1 \leqslant m \leqslant n \\ m \equiv 4 \pmod 8}} \tau(m/4) + \dots$$

$$+ \sum_{\substack{1 \leqslant m \leqslant n \\ m \equiv 2^l \pmod{2^{l+1}}}} \tau(m/2^l)$$

$$= \sum_{r=0}^{(n-1)/2} \tau(2r+1) + \sum_{r=0}^{(n-2)/4} \tau(2r+1) + \dots$$

$$+ \sum_{r=0}^{(n-2^l)/2^{l+1}} \tau(2r+1),$$

and since $l = \left[\dfrac{\log n}{\log 2}\right]$ this is

$$\sum_{m=1}^{n} \gamma(m) = \sum_{t=0}^{(\log n)/\log 2} \sum_{r=0}^{2^{-t}n-1} \tau(2r+1). \tag{2}$$

We estimate the sum

$$\sum_{r=0}^{(w-1)/2} \tau(2r+1)$$

by counting the "odd" lattice points (x, y), i.e., those with both coordinates odd, for which $0 < xy \leqslant w$. (For a full account of this kind of reasoning, see Hardy and Wright, [1], p. 263). We put

$$u = 2[\tfrac{1}{2}w^{\frac{1}{2}}] + 1$$

and obtain

$$\sum_{r=0}^{(w-1)/2} \tau(2r+1) = 2\sum_{z=0}^{(u-1)/2}\left[\frac{1}{2}\left(\frac{w}{2z+1}\right)\right] - \frac{(u-1)^2}{4} + O(1)$$

$$= \frac{1}{4} w \log w + \frac{2C + 2\log 2 - 1}{4} w + O(w^{\frac{1}{2}}).$$

Putting this estimate in (2), we have

$$\sum_{m=1}^{n} \gamma(m) = \sum_{t=0}^{(\log n)/\log 2} \left\{\frac{1}{4}\frac{n \log (n/2^t)}{2^t} + \frac{2C + 2\log 2 - 1}{4}\frac{n}{2^t} + O(2^{-\frac{1}{2}t}n^{\frac{1}{2}})\right\}$$

$$= \frac{n \log n}{2} + \frac{2C + \log 2 - 1}{2} n + O(n^{\frac{1}{2}}),$$

and this completes the proof.

3. A density theorem concerning $\gamma(m)$.

THEOREM 3. *Let ω be a real number, and let $s_n(\omega)$ be the number of positive integers $m \leqslant n$ for which*

$$\gamma(m) \leqslant 2^{\log\log n + \omega(\log\log n)^{\frac{1}{2}} - 1} = f(n, \omega).$$

Then

$$s_n(\omega) \sim nD(\omega).$$

The proof of this is quite similar to that given by Kac [2] in proving that the number of $m \leqslant n$ for which $\tau(m) \leqslant 2f(n, \omega)$ is asymptotic to $nD(\omega)$.

4. Representations in arithmetic progressions. We now turn our attention to $\gamma_1(m, s)$, the number of representations of m of the form

$$m = r + (r + s) + \ldots + \{r + (k - 1)\, s\}. \tag{3}$$

Although it was natural in the case $s = 1$ to restrict attention to positive representations (i.e., with $r > 0$), it turns out in the general case that this condition introduces complications. For this reason we shall consider separately the quantity $\gamma_1(m, s)$ and the quantity $\gamma(m, s)$, the number of positive representations of m in the form (3). In either case it is required that

$$2m = k\{2r + (k - 1)\, s\}. \tag{4}$$

THEOREM 4. $\gamma_1(m, s) = \tau(m)$ *if* $s \equiv 0 (\text{mod } 2)$, *and* $\gamma_1(m, s) = 2\,\tau(\overline{m})$ *if* $s \equiv 1$ (mod 2).

For if s is even, say $s = 2s_1$, then $\gamma_1(m, s)$ is the number of solutions k, r $(k > 0)$ of

$$m = k(r + (k - 1)s_1),$$

and k can clearly be any divisor of m. If s is odd, then k and $2r + (k - 1)s$ are of opposite parity, so that

$$\gamma_1(m, s) = \sum_{k|\overline{m}} 1 + \sum_{2m/k|\overline{m}} 1 = 2\tau(\overline{m}).$$

For example,

$$\begin{aligned} \gamma_1(6,1) = 4: \quad 6 = 1 + 2 + 3 &= (-5) + (-4) + \ldots + 4 + 5 + 6 \\ &= 0 + 1 + 2 + 3; \end{aligned}$$

and

$$\gamma_1(6,2) = 4: \quad 6 = 2 + 4 = 0 + 2 + 4 = (-4) + (-2) + 0 + 2 + 4 + 6.$$

As an immediate consequence of Theorems 2 and 4, and the fact that the average order of $\tau(m)$ is $\log m + 2C - 1 + O(m^{-\frac{1}{2}})$ ([1], *loc. cit.*), we have

THEOREM 5.

$$\frac{1}{n} \sum_{m=1}^{n} \gamma_1(m, s) = \begin{cases} \log n + (2C - 1) + O(n^{-\frac{1}{2}}) & \text{if } s \equiv 0 \pmod 2 \\ \log n + \frac{1}{2}(2C - 1 + \log 2) + O(n^{-\frac{1}{2}}) & \text{if } s \equiv 1 \pmod 2. \end{cases}$$

We now put on the restriction $r > 0$. Then by (4), k must be chosen so that

$$k(k-1)\,s < 2m,$$

or

$$k < \frac{1 + (1 + 8m/s)^{\frac{1}{2}}}{2}.$$

But

$$\left(\frac{2m}{s}\right)^{\frac{1}{2}} < \frac{1 + (1 + 8m/s)^{\frac{1}{2}}}{2} < \left(\frac{2m}{s}\right)^{\frac{1}{2}} + 1,$$

so that we will make an error of not more than 1 if, in computing $\gamma(m, s)$, we count the number of suitable k's which do not exceed $(2m/s)^{\frac{1}{2}}$. Thus by the argument used in proving Theorem 4, we find that if $s = 2s_1$ is even,

$$\gamma(m, s) = \sum_{\substack{k \mid m \\ k \leqslant (2m/s)^{\frac{1}{2}}}} 1 + \epsilon(m, s) = \tau(m, (2m/s)^{\frac{1}{2}}) + \epsilon(m, s),$$

where $\tau(m, x)$ is the number of divisors of m which do not exceed x, and $\epsilon(m, s)$ is either 0 or 1. We put

$$A(n, x) = \sum_{m=1}^{n} \gamma(m, s).$$

Then all those lattice points on the hyperbola $xy = m$ for which $x \leqslant (2m/s)^{\frac{1}{2}}$ are counted in the sum $\sum_1^n \tau(m, (2m/s)^{\frac{1}{2}})$, and by considering all positive m not exceeding n, we see that this sum is exactly the number of lattice points in the region $0 < xy \leqslant n$, $y \geqslant \frac{1}{2} sx$. Counting along vertical lines, we have

$$\begin{aligned}
&\sum_{m=1}^{n} \tau(m, (2m/s)^{\frac{1}{2}}) \\
&\quad = \sum_{x=1}^{(2n/s)^{\frac{1}{2}}} \left\{\left[\frac{n}{x}\right] - \frac{sx}{2} + 1\right\} = n \sum_{x=1}^{(2n/s)^{\frac{1}{2}}} \frac{1}{x} + O(n^{\frac{1}{2}}) - \frac{s}{2} \sum_{x=1}^{(2n/s)^{\frac{1}{2}}} x \;+ \left[\left(\frac{2n}{s}\right)^{\frac{1}{2}}\right] \\
&\quad = n\left\{\log\left(\frac{2n}{s}\right)^{\frac{1}{2}} + C + O(n^{-\frac{1}{2}})\right\} - \frac{s}{4}\left\{\left[\left(\frac{2n}{s}\right)^{\frac{1}{2}}\right]^2 + \left[\left(\frac{2n}{s}\right)^{\frac{1}{2}}\right]\right\} + O(n^{\frac{1}{2}}) \\
&\quad = \frac{n}{2}\log n + n\left(C - \frac{1}{2}\log\frac{s}{2} - \frac{1}{2}\right) + O(n^{\frac{1}{2}}).
\end{aligned}$$

As for the sum $\sum_1^n \epsilon(m, s)$, it does not exceed the number of lattice points on the curves $xy = m \leqslant n$ for which

$$(2m/s)^{\frac{1}{2}} < x \leqslant (2m/s)^{\frac{1}{2}} + 1,$$

i.e., the number of lattice points in the bounded region enclosed by the hyperbolas $xy = n$, $(x-1)^2 s = 2xy$ and the line $l_1\colon y = \frac{1}{2}sx$. But the second of these hyperbolas is asymptotic to the line $l_2\colon y = \frac{1}{2}s(x-1)$. Let the inter-

sections of l_1 and l_2 with $xy = n$ be (x_1, y_1) and (x_2, y_2) respectively, and let the chord joining these points be l_3. Then the sum in question is less than the number of lattice points in the triangle with vertices at $(0, 0)$, (x_1, y_1) and (x_2, y_2), plus the number of lattice points in the triangle with vertices at $(0, 0)$, $(x_2\, y_2)$ and the intersection of l_2 with the x-axis. This follows since l_3 is always above the curve $xy = n$. But it is easy to see that the number of lattice points in a triangle does not exceed one more than the sum of its area and perimeter. Hence

$$\sum_{m=1}^{n} \epsilon(m, s) < \tfrac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ 0 & 0 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} + \tfrac{1}{2} \begin{vmatrix} 1 & 0 & 0 \\ 1 & 0 & 2c_0/s \\ 1 & x_2 & y_2 \end{vmatrix} + (x_1^2 + y_1^2)^{\frac{1}{2}}$$
$$+ 2(x_2^2 + y_2^2)^{\frac{1}{2}} + \left\{(x_2 - x_1)^2 + (y_2 - y_1)^2\right\}^{\frac{1}{2}} + 2c_0/s$$
$$+ \left\{(x_2 - 2c_0/s)^2 + y_2^2\right\}^{\frac{1}{2}}.$$

Substituting the values $x_1 = (2n/s)^{\frac{1}{2}}$, $y_1 = (sn/2)^{\frac{1}{2}}$, $x_2 = \frac{1}{2}\left\{(8n/s + 1)^{\frac{1}{2}} + 1\right\}$ $y_2 = n/x_2$, it is easily verified that this upper bound is $O(n^{\frac{1}{2}})$.

We have thus shown that in case s is even,

$$A(n, s) = \frac{n}{2} \log n + \frac{n}{2}\left(2C - \log \frac{s}{2} - 1\right) + O(n^{\frac{1}{2}}). \tag{5}$$

On the other hand, if $s \equiv 1 \pmod 2$, then in (4) either k is even, in which case it contains the highest power 2^e of 2 which divides $2m$ and is such that r is positive, or k is odd, with r again positive. Hence

$$\gamma(m, s) = \sum_{\substack{k|\overline{m} \\ k \leq (2m/s)^{\frac{1}{2}}}} 1 + \sum_{\substack{k_1|\overline{m} \\ 2^e k_1 \leq (2m/s)^{\frac{1}{2}}}} 1 + \epsilon(m, s)$$

$$= \tau(\overline{m}, (2^e\overline{m}/s)^{\frac{1}{2}}) + \tau(\overline{m}, (2^{-e}\overline{m}/s)^{\frac{1}{2}}) + \epsilon(m, s),$$

where $\epsilon(m, s)$, as before, is the error made in assuming that for r to be positive k must not exceed $(2m/s)^{\frac{1}{2}}$, rather than the actual upper bound. Since the bound for $\sum_1^n \epsilon(m, s)$ which we just computed did not depend on the parity of s, it holds also for odd s:

$$\sum_{m=1}^{n} \epsilon(m, s) = O(n^{\frac{1}{2}}). \tag{6}$$

We have

$$A(n, s) = \sum_{m=1}^{n} \tau(\overline{m}, (2^e\overline{m}/s)^{\frac{1}{2}}) + \sum_{m=1}^{n} \tau(\overline{m}, (2^{-e}\overline{m}/s)^{\frac{1}{2}}) + \sum_{m=1}^{n} \epsilon(m, s)$$
$$= A_1 + A_2 + A_3,$$

say. Summing over m's containing the same power of 2, we get

$$A_1 = \sum_{\lambda=1}^{(\log n)/\log 2} \sum_{r=1}^{2^{-\lambda}n - \frac{1}{2}} \tau(2r + 1, \{2^\lambda(2r + 1)/s\}^{\frac{1}{2}}).$$

The sum

$$\sum_{r=0}^{(z-1)/2} \tau(2r+1, c^{\frac{1}{2}}(2r+1)^{\frac{1}{2}})$$

is the number of lattice points on the hyperbolas

$$xy = 2r+1, \qquad r = 0, 1, \ldots, \tfrac{1}{2}(z-1)$$

for which $x \leqslant c^{\frac{1}{2}}(2r+1)^{\frac{1}{2}}$, i.e., for which $x \leqslant cy$. This is the number of odd lattice points in this region, which is

$$\sum_{x=0}^{t} \left\{\left[\frac{1}{2}\left(\frac{z}{2x+1}+1\right)\right] - \left[\frac{1}{2}\left(\frac{2x+1}{c}+1\right)\right] + \delta(x)\right\},$$

where $\delta(x)$ is 0 or 1 and

$$t = \left[\frac{1}{2}\left\{c^{\frac{1}{2}}\left(2\left[\frac{z-1}{2}\right]+1\right)^{\frac{1}{2}} - 1\right\}\right] \sim \frac{(cz)^{\frac{1}{2}}}{2}.$$

But this sum is equal to

$$\frac{z}{2}\sum_{x=0}^{t}\frac{1}{2x+1} + O(t) - \frac{1}{2c}\sum_{x=0}^{t}(2x+1) + O(t)$$

$$= \frac{z}{4}\log\left\{c\left(2\left[\frac{z-1}{2}\right]+1\right)\right\}^{\frac{1}{2}} + \frac{z}{4}(C+\log 2) + O(z^{\frac{1}{2}}) - \frac{t^2}{2c} + O(t),$$

so that

$$\sum_{r=0}^{(z-1)/2} \tau(2r+1, c^{\frac{1}{2}}(2r+1)^{\frac{1}{2}}) = \frac{z\log z}{8} + \frac{z}{4}(C+\log 2+\log c^{\frac{1}{2}}) - \frac{z}{8} + O(z^{\frac{1}{2}}). \tag{7}$$

Hence

$$A_1 = \sum_{\lambda=1}^{(\log n)/\log 2}\left\{\frac{n}{2^{\lambda-1}}\frac{1}{8}\log\frac{n}{2^{\lambda-1}} + \frac{n}{2^{\lambda}}\frac{1}{4}\left(C+\log 2+\log\frac{2^{\frac{1}{2}\lambda}}{s^{\frac{1}{2}}}\right) - \frac{n}{8.2^{\lambda-1}} + O\left(\frac{n^{\frac{1}{2}}}{2^{\frac{1}{2}(\lambda-1)}}\right)\right\}$$

$$= \frac{n\log n}{8}\sum_{\lambda=1}^{\log n}\frac{1}{2^{\lambda-1}} - \frac{n}{8}\log 2\sum_{\lambda=1}^{\log n}\frac{\lambda-1}{2^{\lambda-1}} + \frac{n}{4}(C+\log 2)\sum_{\lambda=1}^{\log n}\frac{1}{2^{\lambda-1}}$$

$$+ \frac{n\log 2}{8}\sum_{\lambda=1}^{\log n}\frac{\lambda}{2^{\lambda-1}} - \frac{n\log s}{8}\sum_{\lambda=1}^{\log n}\frac{1}{2^{\lambda-1}} - \frac{n}{8}\sum_{\lambda=1}^{\log n}\frac{1}{2^{\lambda-1}} + O(n^{\frac{1}{2}})$$

$$= \frac{n\log n}{4} + n\left(\frac{C}{2}+\frac{\log 2}{2}-\frac{\log s}{4}\right) - \frac{n}{4} - \frac{n\log 2}{4}$$

$$+ \frac{n\log 2}{4} + O(n^{\frac{1}{2}}),$$

and finally

$$A_1 = \frac{n \log n}{4} + n\left(\frac{C + \log 2}{2} - \frac{1}{4} - \frac{\log s}{4}\right) + O(n^{\frac{1}{2}}). \tag{8}$$

Turning now to A_2, we have

$$A_2 = \sum_{\lambda=1}^{(\log n)/\log 2} \sum_{r=0}^{2^{-\lambda}n-\frac{1}{2}} \tau\left(2r+1, \left(\frac{2r+1}{2^{\lambda}s}\right)^{\frac{1}{2}}\right),$$

and using (7) with $z = n/2^{\lambda-1}$, $c = s/2^{\lambda}$, we have

$$\begin{aligned} A_2 = \sum_{\lambda=1}^{(\log n)/\log 2} & \left\{\frac{n}{8.2^{\lambda-1}} \log \frac{n}{2^{\lambda-1}} + \frac{n}{4.2^{\lambda-1}} (C + \log 2 + \log (2^{\lambda}s)^{-\frac{1}{2}}) \right. \\ & \left. - \frac{n}{4.2^{\lambda-1}} + O\left(\frac{n^{\frac{1}{2}}}{2^{\frac{1}{2}(\lambda-1)}}\right)\right\} \\ = \frac{n \log n}{4} & + n\left(\frac{C}{2} - \frac{1}{4} - \frac{\log s}{4}\right) + O\,(n^{\frac{1}{2}}). \end{aligned}$$

Combining this with (5), (6) and (8), we have

THEOREM 6. *For every* s,

$$\frac{1}{n} \sum_{m=1}^{n} \gamma(m, s) = \frac{1}{2} \log n + \left(C - \frac{1}{2} \log \frac{s}{2} - \frac{1}{2}\right) + O\,(n^{-\frac{1}{2}}).$$

Theorem 2 is, of course, the special case of Theorem 6 with $s = 1$.

REFERENCES

[1] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers* (Oxford, 1945).

[2] M. Kac, *Note on the distribution of values of the arithmetic function* $d(m)$, Bulletin Amer. Math. Soc., vol. 47 (1941), 815-817.

University of Michigan

THE ITERATION OF CERTAIN ARITHMETIC FUNCTIONS

IVAN NIVEN

1. Introduction. For $n \geqslant 3$ define $C(n)$ to be the integer j such that $\phi^{(j)}(n) = 2$, where $\phi^{(j)}(n)$ denotes the jth iterate of the Euler ϕ-function. Define $C(1) = C(2) = 0$. This function has been studied by S. S. Pillai [1], with the notation $R(n)$ for $1 + C(n)$ if $n \geqslant 2$, and $R(1) = 0$. H. Shapiro [2] has also investigated this function, proving the basic relations

$$C(ab) = C(a) + C(b) \text{ or } C(ab) = C(a) + C(b) + 1, \tag{1}$$

the second equation holding when a and b are both even, otherwise the first.

It was suggested to the writer by Morgan Ward that a function analogous to $C(n)$ can be obtained by iteration of $\lambda(n)$, the least positive exponent so that

$$a^{\lambda(n)} \equiv 1 \pmod{n} \tag{2}$$

for every a which is prime to n. Thus for $n \geqslant 1$ we define $g(n)$ as the least positive integer j such that $\lambda^{(j)}(n) = 1$, where $\lambda^{(j)}(n)$ is the jth iterate of the λ-function. We now prove the following results.

THEOREM 1. *If* $(a, b) = 1$, *then* $g(ab) = \max\{g(a), g(b)\}$.

THEOREM 2. *For* $n \geqslant 1$, $g(2^{2n}) = g(2^{2n+1}) = n + 1$, $g(p^n) = n - 1 + g(p)$ *where* p *is any odd prime.*

The method of deriving functions $C(n)$ and $g(n)$ from $\phi(n)$ and $\lambda(n)$ can be generalized to obtaining $F(n)$ from any $f(n)$ which has the property $f(n) < n$ for $n > k$, where k is a constant. It might be expected that $F(n)$ would have a property similar to (1) whenever $f(n)$ was multiplicative, that is, whenever $f(ab) = f(a)f(b)$ for relatively prime a and b. That this is not so can be seen readily by taking $f(n)$ to be the number of divisors of n. Similarly, Theorem 1 is not implied merely by the functional relation of $\lambda(n)$, namely

$$\lambda(ab) = \text{l.c.m.}\ \{\lambda(a), \lambda(b)\} \text{ whenever } (a, b) = 1. \tag{3}$$

In §2 we shall prove Theorems 1 and 2, and the next two theorems in §3 and §4.

THEOREM 3. $\limsup\{C(n + 1) - C(n)\} = \limsup\{g(n + 1) - g(n)\}$
$= \limsup\{C(n) - g(n)\} = \infty$,

THEOREM 4. $\liminf\{C(n + 1) - C(n)\} = \liminf\{g(n+1) - g(n)\} = -\infty$,
$\liminf\{C(n) - g(n)\} = -1$.

Received September 4, 1949.

2. The fundamental results for $g(n)$. It is known that for any odd prime p,

(4) $$\lambda(p^n) = \phi(p^n) = p^{n-1}(p-1),$$

(5) $$\lambda(2^n) = \tfrac{1}{2}\phi(2^n) = 2^{n-2} \text{ for } n \geqslant 3; \quad \lambda(4) = 2; \quad \lambda(2) = 1.$$

Together with (3), these imply

(6) $$\lambda(m) | \lambda(n) \text{ whenever } m|n.$$

Now Theorem 1 clearly holds if $a = 1$, and we use mathematical induction, assuming the result for $g(n)$ with $n < ab$. Ignore the trivial cases where $g(a) = g(b) = 1$, or where $a = 1$ or $b = 1$. We have

(7) $$g(n) = 1 + g(\lambda(n)) \quad \text{for } n > 2,$$

and so

(8) $$g(ab) = 1 + g(\lambda(ab)) = 1 + g\{\text{l.c.m. } (\lambda(a), \lambda(b))\}.$$

Also $\lambda(a) < a$, $\lambda(b) < b$, so that $ab >$ l.c.m. $(\lambda(a), \lambda(b)) = p_1^{a_1}p_2^{a_2} \ldots p_r^{a_r}$, these primes being arranged so that $g(p_1^{a_1}) \geqslant g(p_i^{a_i})$ $(i = 2, 3, \ldots, r)$. Thus by (6) and the induction hypothesis, (8) becomes $g(ab) = 1 + g(p_1^{a_1})$. Without loss of generality we may assume that $p_1^{a_1}$ is a divisor of $\lambda(a)$, so that $g(p_1^{a_1}) = g(\lambda(a)) \geqslant g(\lambda(b))$, whence $g(a) \geqslant g(b)$ by (7). Hence we have $g(ab) = 1 + g(\lambda(a)) = g(a) = \max\{g(a), g(b)\}$.

To prove Theorem 2, we note that the first part is established by (5). And the second part can be obtained by use of mathematical induction, (4), (7) and Theorem 1. Thus for $n \geqslant 2$,

$$g(p^n) = 1 + g(\lambda(p^n)) = 1 + g\{p^{n-1}(p-1)\} = 1 + g(p^{n-1}).$$

3. Proof of Theorem 3. We shall in this and the following section make use of two results of Pillai [1, Theorems 1 and 3] which can be summarized thus:

(9) $$[\log_2 n] \geqslant C(n) \geqslant \log_3 n/2.$$

Since $4^{3^k} - 1$ or $(1+3)^{3^k} - 1$ is divisible by 3^k we can write, using (9) and (1) and the fact that $C(3^k) = k$,

$$\begin{aligned} C(4^{3^k} - 1) &= C(3^k) + C\{(4^{3^k} - 1)/3^k\} \\ &< k + \log_2 \{(4^{3^k} - 1)/3^k\} \\ &< k + 2 \cdot 3^k - k \log_2 3. \end{aligned}$$

Also $C(2^j) = j - 1$ and so we have

$$C(4^{3^k}) - C(4^{3^k} - 1) > 2 \cdot 3^k - 1 - k - 2 \cdot 3^k + k \log_2 3 = k \log_2 (3/2) - 1.$$

This establishes the first part of Theorem 3.

By (4) and (5) we have $g(n) \leqslant C(n) + 1$, and so (9) implies

(10) $$g(n) \leqslant 1 + [\log_2 n].$$

Now $(3^k + 1, 3^k - 1) = 2$ and we apply Theorem 1 to get

$$g(3^{2k}-1) \leqslant 1 + \max\{g(3^k+1), g(3^k-1)\}$$
$$\leqslant 2 + \log_2(3^k+1) < 3 + k\log_2 3.$$

From this it follows that

(11) $$g(3^{2k}) - g(3^{2k}-1) > 2k+1-3-k\log_2 3,$$

which proves the second part of Theorem 3.

The last part of Theorem 3 can be obtained by taking n to be the product of the first k primes, and using (1), Theorem 1, (9) and (10).

4. Proof of Theorem 4. By (9) we see that

(12) $$C(3^j+1) \geqslant j.$$

Next we prove that

(13) $$C(3^{2^k}-1) \geqslant 2^k + k - 1$$

by mathematical induction. Using (1) and (12) we have

$$C(3^{2^k}-1) = C(3^{2^{k-1}}+1) + C(3^{2^{k-1}}-1) + 1$$
$$\geqslant 2^{k-1} + 2^{k-1} + k - 2 + 1.$$

Having proved (13), we see that it implies

$$C(3^{2^k}) - C(3^{2^k}-1) \leqslant 2^k - 2^k - k + 1 = -k+1,$$

which establishes the first part of Theorem 4.

We now discuss $g(n+1) - g(n)$ with $n = (3^{2k}-1)^2$, k odd. Thus $3^{2k} \equiv 9 \pmod{16}$ and $3^{2k} \equiv -1 \pmod 5$, so that $3 \nmid n$, $5 \nmid n$, $2^6 \mid n$, $2^7 \nmid n$. So for large k we have $g(n) = g(p^{2j})$ where p is some odd prime > 5 and $p^j < 3^k+1$ so that $j < 1 + k\log_p 3$. Using (10) we have

(14) $$g(n) = g(p^{2j}) = 2j - 1 + g(p) < 2j + \log_2 p < 2 + 2k\log_p 3 + \log_2 p.$$

Considering the last expression as a function of a continuous variable p on the range $(7, 3^k)$, with k constant, we see that it is a maximum for $p = 3^k$, so that (14) implies $g(n) < 4 + k\log_2 3$. Hence we have

$$g(n+1) - g(n) > g\{3^{2k}(3^{2k}-2)\} - 4 - k\log_2 3$$
$$\geqslant g(3^{2k}) - 4 - k\log_2 3$$
$$= 2k + 1 - 4 + k\log_2 3.$$

This proves the second part of Theorem 4, and the final part is a consequence of the two results $g(n) \geqslant C(n) + 1$ and $g(3^k) = k + 1 = 1 + C(3^k)$.

References

[1] S. S. Pillai, *On a function connected with* $\phi(n)$, Bull. Amer. Math. Soc., vol. 35 (1929), 837-841.

[2] Harold Shapiro, *An arithmetic function arising from the* ϕ *function*, Amer. Math. Monthly, vol. 50 (1943), 18-30.

University of Oregon

ITERATES OF FRACTIONAL ORDER

RUFUS ISAACS

1. Introduction. The body of this paper is a complete answer to the following question:

Let E be any space whatever. $g(x)$ is a function[1] mapping E into E. When does there exist a function $f(x)$, of the same type, such that

$$f(f(x)) = g(x) \qquad (x \in E)? \tag{1}$$

This problem typifies the general one of iteration. Let $g^k(x)$ be the kth order iterate of g [i.e. $g^0(x) = x$, $g^{k+1}(x) = g(g^k(x))$]. The iteration problem is that of attaching a consistent meaning to this expression for fractional k (in the sense of preserving the additive law of exponents). An f satisfying (1) is thus $g^{1/2}(x)$. By ideas similar to those discussed herein, we can find the most general $g^{1/m}$ and then by iterating it, the most general iterate of any rational order. Without introducing continuity, this is as far as it is possible to go. We confine ourselves to the case of $k = 1/2$ to avoid oppressive detail; the generalization to $k = 1/m$ is indicated later.

The iteration problem has received attention for many years, alone or as part of another topic (functional equations, fractional derivatives, the tri-operational algebra of Menger [1], etc.). Some of these applications require subsidiary conditions on the functions (continuity, differentiability, etc.). We deal with the general problem without such side conditions; thus our work might be called combinatorial. The problem with a side condition such as continuity appears highly interesting.

In all the literature we have encountered, the general problem is approached in but one way—through the Abel function. The idea here is to ascertain a numerically valued function ϕ on E satisfying

$$\phi(g(x)) = \phi(x) + 1.$$

Then iterates of all orders are obtained at once by

$$g^k(x) = \phi^{-1}(\phi(x) + k).$$

We show later that in a widespread class of cases, a ϕ does not exist. Even when it does, its inverse may not exist. Yet iterates of some or all fractional orders may exist. The non-existence of ϕ may hold even when we have continuity with respect to both x and k, as we shall show below.

Received April 12, 1949.

[1] If g is not defined for all of E, it suffices that our later criterion hold for some extension of g which is. If the range and domain of g are distinct we can thus take E to be their union.

For the Abel function approach in the complex number domain see the papers by Schwarzschild, Chayoth, and Koenigs [2]. For the real domain, Lyche [3] gives existence conditions for ϕ and continuous ϕ by methods somewhat akin to ours. Bödewadt [4] treats the case of a fully differentiable ϕ (real domain). Hadamard [5] summarizes two recent contributions.

Our interest in this question arose from the following problem propounded by Menger. Let E be R_1 and $g(x) = a + bx$. There is obviously a linear solution to (1) when $b \geqslant 0$, namely $f(x) = a/(1 + b^{\frac{1}{2}}) + b^{\frac{1}{2}}x$. Do solutions exist when $b < 0$? The question is answered below.

The text will be clearer if we outline our method first.

An *orbit* (defined precisely later) is a subset of E whose elements are linked by the operation g. We can represent one graphically as in Figure 1 where the dots represent elements of E and the arrows show the course of g.

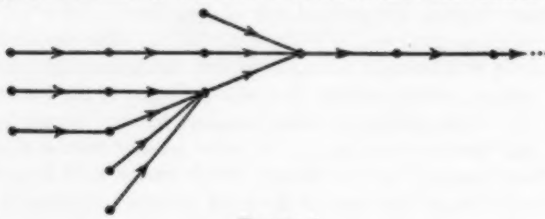

FIGURE 1

The present idea consists of constructing the orbits of f from those of g. Thus in Figure 2 we see two orbits with respect to g united so as to give one with respect to f. The dashed arrows show the course of f; the truth of (1) may be verified by noting that following two consecutive dashed arrows is equivalent to following one solid one.

This kind of construction can sometimes be carried out utilizing only a single g-orbit as in Figure 3.

We will show (Theorem 1) that these two instances typify the most general situation possible. The problem then reduces essentially to two questions: *When can two distinct g-orbits be "mated"* (*as in Figure* 2) *to produce one for f*? *When can a single g-orbit also be an f-orbit* (*as in Figure* 3)? which are answered by Theorems 3 and 4.

2. Orbits. Consider the following relation between the members x,y of E: There exist non-negative integers m, n such that

$$g^m(x) = g^n(y). \tag{2}$$

This relation is rst[2]; the classes into which it divides E, in the customary way, are called *orbits*.[3] The orbit containing x will be denoted by $L[x;g]$.

[2]Reflexive, symmetric and transitive.

[3]This concept appears in Lyche [3], where he attributes it to a suggestion of Kuratowski. He uses the term *class*; in a previous abstract of this work we used *linkage*. The term *orbit* appears in Whyburn [6].

A set

(3) $$x_1, \ldots, x_n$$

such that $g(x_1) = x_2, \ldots, g(x_n) = x_1$ will be called a *cycle* (*n-cycle*).

LEMMA 1. *An orbit contains at most one cycle.*

For let x and y be elements which belong to the same orbit but to two distinct cycles; (2) holds. The element on both its sides belongs to both cycles. The cycles, having a common element, are identical.

FIGURE 2

An orbit containing a cycle (of n elements) will be called *cyclic* (*n-cyclic*). Let C be the cycle of a cyclic orbit L. An element x_0 will be called a *leader* if

$$x_0 \in L - C, \quad g(x_0) \in C.$$

For a particular leader x_0 the subset of all y of L such that for some non-negative integer n

(4) $$g^n(y) = x_0$$

will be called a *branch* or more precisely a *branch from* $g(x_0)$.

LEMMA 2. *The branches constitute an aliquot, disjoint subdivision of* $L - C$. *For each* y, *the* n *of* (4) *is unique.*

Let $y \in L - C$; (2) holds for y and any x which $\in C$. Let n_1 be the smallest n such that $g^n(y) \in C$; $n_1 > 0$. Put $n = n_1 - 1$. Then $g^n(y) = x_0$ is a leader. This x_0 and n are unique, for suppose the existence of a second pair, i.e.

$$g^{n'}(y) = x'_0$$

and say $n \geqslant n'$. Then $x_0 = g^n(y) = g^{n-n'}(x'_0)$. Now $n - n' > 0$ is impossible as this implies $x_0 \in C$. Then $n = n'$, $x_0 = x'_0$.

Any union of branches from the same $z \in C$ is called a *branch cluster* (*from* z).

The subset of all y of a branch B for which the n of (4) is even (odd) is called the *even* (*odd*) *part of* B.

The following two operations concern only the structural properties of orbits, i.e. those invariant under isomorphisms (the term is used in the expected

sense of a biunique, g-preserving correspondence). In other words, we admit orbits whose elements are abstract.

Consider the subsets X of an orbit L which are inverse images of single elements of E under g. (X is the set of $x \in L$ such that $g(x) = y$ for some fixed $y \in L$.) Divide each such X into a system of aliquot, disjoint subsets X_a. Identify the elements of each X_a into a single element, thus obtaining a new orbit L'. For L', g is defined by $g(X_a) = g(x)$ where $x \in X_a$; if $g(y) = x \in X_a$ then for L', $g(y) = X_a$. L' will be called a *contraction* of L. For

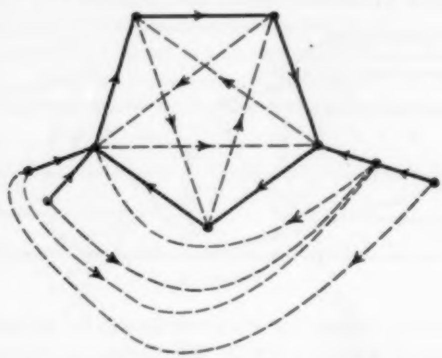

FIGURE 3

a cyclic orbit we may apply the idea to its branches. We include the possibility of contracting a branch cluster into a branch by identifying all its leaders.

By a *curtailment* of an orbit or branch L is meant the new orbit or branch arising when some of the elements x of L for which there is no y such that $g(y) = x$ are removed from L. (For the unremoved elements, g is unchanged.)

3. The existence conditions.

THEOREM 1. *If f, g satisfy* (1), *each orbit with respect to f is the union of two (possibly identical) orbits with respect to g. More precisely*:

$$L(x;f) = L(x;g) \cup L(f(x);g). \qquad (x \in E) \tag{5}$$

Let $y \in L(x;f)$. Then, for suitable m, n,

$$f^m(y) = f^n(x) \tag{6}$$

and also

$$f^{m+1}(y) = f^{n+1}(x). \tag{7}$$

One of (6), (7) has an even superscript on the left; let it be

$$f^{2k}(y) = f^{2j+\epsilon}(x), \qquad \epsilon = 0 \text{ or } 1$$

which can also be written

$$g^k(y) = g^j(f^i(x))$$

so that $y \in$ the right side of (5).

On the other hand, if $y \in L(x; g)$ or $L(f(x); g)$, (2) can be written

$$f^{2n}(y) = f^{2m}(x) \text{ or } f^{2m+1}(x).$$

Two distinct orbits capable of being paired together in the manner mentioned in Theorem 1 are said to be *mateable*. An orbit capable of being paired with itself will be said to be *self-mateable*.

The existence criterion for f is now clear.

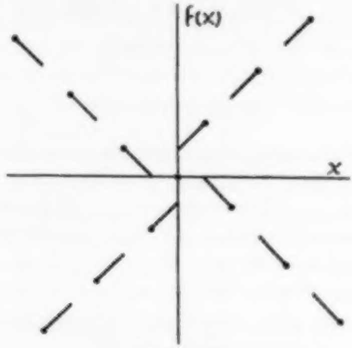

FIGURE 4

THEOREM 2. *A necessary and sufficient condition for f to exist is that the set of orbits with respect to g can be divided into three aliquot, disjoint subsets, such that two can be put into a biunique correspondence with mateable correspondents, while the third consists of self-mateable orbits.*

It remains to find criteria for mateability and self-mateability.

THEOREM 3. *A necessary and sufficient condition for two distinct orbits to be mateable is that a contraction of one be isomorphic to a curtailment of the other.*

Sufficiency. Let L_1, L_2 be orbits such that a contraction of L_1 is isomorphic to a curtailment L'_2 of L_2. If $x \in L_1$ then a subset of L_1, containing x, is paired by the isomorphism to $y \in L'_2$; define $f(x) = y$ and $f(y) = g(x)$. If $y \in L_2 - L'_2$, let $f(y)$ be any element of L_1 which is mapped by f into $g(y)$ (possible, as $g(y) \in L'_2$). The so-defined f satisfies (1).

Necessity. Let L_1, L_2 be the orbits. Identifying the x of L_1 for which $f(x)$ is the same element of L_2 gives a contraction L'_1 of L_1. Then f establishes an isomorphism between L'_1 and a subset of L_2. The excluded elements of L_2 may be removed by a curtailment.

Since the presence of an n-cycle is invariant under contractions and curtailments we have

COROLLARY. *n-cyclic orbits are mateable only with n-cyclic orbits.*

THEOREM 4. *A necessary and sufficient condition for an orbit L to be self-mateable is*

1) *L is n-cyclic with n odd. Let $n = 2k + 1$.*

2) *The branches of L are disjointedly the union of a set of branches S and a set of branch clusters $\overline{S}$. The S and $\overline{S}$ are in a biunique correspondence such that if $B \in S$ and $\overline{B} \in \overline{S}$ correspond, then a contraction of $\overline{B}$ is isomorphic to a curtailment of B and*[4] *if B is from z, $\overline{B}$ is from $g^k(z)$.*

Necessity. Let $x \in L$. As $f(x) \in L$, for suitable p, q,

$$g^p(x) = g^q(f(x))$$

or

$$f^{2p}(x) = f^{2q+1}(x). \tag{8}$$

As the two superscripts are distinct, familiar reasoning shows that for some j, $f^j(x)$ belongs to a cycle. Let it be C of order n. Let (3) be its elements so numbered that[5] $f(x_j) = x_{j+1}$. Then[5] $g(x_j) = x_{j+2}$. If n were even, the subsets of (3) with odd and even subscripts would each constitute a distinct cycle of L with respect to g. Put $n = 2k + 1$.

If $x \in C$, then $f(x) = f^{2k+2}(x) = g^{k+1}(x)$.

Now let B' be a branch with respect to f; x_0, its leader; B and $\overline{B}$, its even and odd parts. As $x_0 \in B$, B is not vacuous.

Letting $y \in B$, we must have for some $m \geqslant 0$

$$f^{2m}(y) = x_0 = g^m(y).$$

As x_0 is a leader with respect to g also, we see that, in regard to g, B is a branch from $g(x_0) = z$.

Similarly, if $y \in \overline{B}$,

$$f^{2m+1}(y) = x_0 = g^m(f(y))$$

which implies

$$g^{m+1}(y) = f(x_0) \in C.$$

Thus, in regard to g, $\overline{B}$ is a branch cluster from

$$f(x_0) = f^{2k+1}(f(x_0)) = g^{k+1}(x_0) = g^k(z).$$

Thus we have supplied the correspondence mentioned in 2). That a contraction of $\overline{B}$ is isomorphic to a curtailment of B follows as in the proof of Theorem 3.

---

[4]We admit vacuous branch clusters, but not vacuous branches.

[5]Reckoned mod n.

Sufficiency. If x is in the cycle of the given orbit we define:

$$f(x) = g^{k+1}(x).$$

Now let B and $\bar{B}$ be as in 2). An f can be defined for their members as in the proof of Theorem 3, with evident modifications.

4. Inadequacy of the Abel function method. Lyche has shown that (in the case of functions of a real variable, but the result is true generally):

A necessary and sufficient condition for the Abel function to exist is that for no positive integer n and $x \in E$ is $g^n(x) = x$.

In other words, the condition is that there be no cyclic orbits. Our conditions show that f may exist in the contrary case. For example, let the orbit diagrammed in Figure 3 comprise the entire space E.

The truth of a fixed point theorem is equivalent to the existence of a 1-orbit. Thus the non-existence of the Abel function is not uncommon.

Now let E be the set of all non-negative numbers and $g(x) = x^2$. If we define $g^k(x)$ to be x^{2^k} we have a consistent iterate for each real k. Yet ϕ does not exist as 0 and 1 each belong to a 1-cycle.

We can easily construct the Abel function using the diagrams of non-cyclic orbits. In Figure 1, say, assign a real number to each vertical bank of dots in such a way that these numbers increase by unity as we proceed to the right. Doing this for each orbit (assumed non-cyclic), we obtain the most general Abel function. The truth of Lyche's theorem now becomes apparent.

For an Abel function to have an inverse it is clearly necessary that each vertical bank contain at most one dot. Further, the numbers must be assigned so as to avoid duplication of values on different orbits. If an Abel function is to be usuable for constructing iterates of all real orders, there must be a large enough number of orbits for each real number to occur once among its function values.

5. Examples: *The Menger Problem.* Let E be R_1 and $g(x) = a + bx$. If $b < 0$, our technique enables us still to construct solutions of (1), but they will never be continuous.

To illustrate, we take the case: $g(x) = -x$. Here, the orbit containing 0 is a 1-cycle. All other orbits are 2-cycles containing x and $-x (x \neq 0)$; there is thus exactly one containing a given positive number. The former can and must be self-mated; the latter are mateable in pairs.

To construct an example we must first divide the set of positive numbers into two parts in biunique correspondence. Taking these parts, say, to be the alternate intervals $(n, n+1]$ and for the correspondence, using an obvious linear mapping, we are led to a function whose graph is sketched in Figure 4. (The heavy dots on the ends of the segments indicate that these end points are included.)

The problem has continuous solutions if we work in the complex domain. On the other hand there exist analytic g such that (1) has a continuous solution

in the real domain, but none at all in the complex domain. Such is $g(x) = x^2$. In the real domain take $f(x) = |x|^{2^{1/2}}$. In the complex domain no f exists as there is but one 2-cycle (namely, the complex cube roots of unity.)

Iterates of order $1/m$. It is not hard to generalize from (1) to

$$f^m(x) = g(x).$$

We state without proof the partial result:

Each orbit L_0 *with respect to* f *is the union of orbits* $L_1, \ldots, L_p$ *with respect to* g *and* p *is a divisor of* m. *If* $p < m$, L_0 *is cyclic. When* L_0 *is cyclic of order* n, $L_1, \ldots, L_p$ *are all cyclic of order* n/p, *and*

$$p = (m, n).$$

The oddness of n in Theorem 4 follows from the special instance of this last equation: $p = 1$, $m = 2$.

References

[1] Menger, K., *Tri-operational algebra*. Rep. Math. Coll., University Notre Dame, issue 5-6, (1944).

[2] Schwarzschild, *Über eine Interpolationsaufgabe der Atkinometrie*. Astron. Nachr., 172, (1906), 65-74, F. d. M37, 963.

W. Chayoth, *Stetige Lösungen gewisser Funktional Gleichungen*, Monatshefte für Math., vol. 39, (1932), 279-288.

Koenigs, G., *Recherches sur les intégrales de certaines équations fonctionelles*, Ann. de l'Ecole Norm. Sup. (1882), 1-41.

——— *Sur les intégrals de certaines équations fonctionelles*, CR IC 1016-1017 (1882).

[3] Tambs Lyche, R., *Sur l'équation fonctionelle d'Abel*, Fund. Math., vol. 5, (1924), 331-333.

[4] Bödewadt, U. T., *Zur Iteration reeller Funktionen*, Math. Zeitschrift, vol. 49, No. 3, (1944).

[5] Hadamard, J., *Two works on iteration and related questions*, Bull. Amer. Math. Soc., vol. 50, (1944), 67-75.

[6] Whyburn, G. T., *Analytic topology*, Amer. Math. Soc., Coll. Publ., No. XXVIII.

University of Notre Dame
and Rand Corporation, Santa Monica

LATTICES WITH A GIVEN ABSTRACT GROUP OF AUTOMORPHISMS

ROBERT FRUCHT

THE problem of finding a lattice[1] with a given abstract group of automorphisms has been solved by Garrett Birkhoff[2] who proved that for any group of order g there exists a distributive lattice with at most 2^{g^2+g} elements. That this number can be somewhat reduced by modifications of Birkhoff's original procedure has already been shown by the author[3]; it turns out, however, that it remains rather high for finite groups of relatively low order.

The purpose of the present paper is to show that a lattice with fewer elements can be found by a completely different method; in general, however, this lattice will not be distributive. Indeed we shall prove (see Theorem 2 below) that for any group of finite order g which can be generated by n of its elements a lattice can be found with at most $5(n + 2)\,g + 2$ elements. (To obtain an upper bound independent of n it suffices to recall that always $n \leqslant \dfrac{\log g}{\log 2}$.)

Since our method of finding a lattice with a given group of automorphisms is rather closely related to some theorems on graphs and their groups, we begin by recalling the definitions of these two notions.

By a *graph* we mean a finite set of elements called vertices some of which are joined by edges (or arcs), but so that two vertices are never joined by more than one edge; also the case of isolated vertices (which are not endpoints of any edge) will be excluded. If in a graph with q vertices $P_1, P_2, \ldots, P_q$ we define incidence-numbers I_{P_i,P_k} $(i \neq k)$ by

$$I_{P_i,P_k} = I_{P_k,P_i} = \begin{cases} 0, & \text{if } P_i \text{ and } P_k \text{ are not joined by an edge,} \\ 1, & \text{if } P_i \text{ and } P_k \text{ are joined by an edge,} \end{cases}$$

then the graph itself may also be characterized by the following quadratic form in q indeterminates $x_1, x_2, \ldots, x_q$:

$$F(x_1, x_2, \ldots, x_q) = \sum_{i<k} I_{P_i,P_k} x_i x_k.$$

Received September 26, 1949.

[1]For the definition of lattice and other basic notions of lattice theory, see Garrett Birkhoff's *Lattice Theory* (1st ed., New York, 1940).

[2]Garrett Birkhoff, *Sobre los grupos de automorfismos.* Revista de la Union Matematica Argentina, vol. 11 (1946), pp. 155-157.

[3]R. Frucht, *Sobre la construccion de sistemas parcialmente ordenados con grupo de automorfismos dado.* Revista de la Union Matematica Argentina, vol. 13 (1948), pp. 12-18. See also: *On the construction of partially ordered systems with a given group of automorphisms.* Amer. J. Math., vol. 72 (1950), pp. 195-199.

The *group* (of automorphisms) of the graph then consists of those permutations of $x_1, x_2, \ldots, x_q$ which leave the quadratic form $F(x_1, x_2, \ldots, x_q)$ unaltered; it is obvious that the corresponding permutations of the vertices $P_1, P_2, \ldots, P_q$ of the graph represent all the possible mappings of the graph into itself which preserve incidence-relations.

The connexion between lattices and graphs is given by the following general theorem.

THEOREM 1. *Given any graph (in the sense defined above) with q vertices and p edges, there is always a lattice with $p + q + 2$ elements such that the group of automorphisms of the lattice is simply isomorphic to that of the graph.*

Proof. Let $P_1, P_2, \ldots, P_q$ be the vertices of the given graph G, and let $a_1, a_2, \ldots, a_p$, be its edges. A partially ordered system S with $p + q + 2$ elements $I, A_1, A_2, \ldots, A_p, B_1, B_2, \ldots, B_q, O$ may then be defined by the following order-relations:

(1) $I > A_i > O$ (for $i = 1, 2, \ldots, p$),

(2) $I > B_j > O$ (for $j = 1, 2, \ldots, q$),

(3) $A_i > B_j$ if, and only if, the vertex P_j is in G one of the endpoints of the edge[4] a_i.

This system S is a lattice, as it is evident that any two of its elements have always a greatest lower bound or meet (symbol: $\cap$) and a lowest upper bound or join (symbol: $\cup$); e.g. it is obvious that

$$A_i \cup A_k = I \text{ for any } i \neq k,$$

and that

$$A_i \cap A_k = \begin{cases} O, & \text{if in } G \text{ the edges } a_i \text{ and } a_k \text{ have no common endpoint,} \\ B_j, & \text{if in } G \text{ the edges } a_i \text{ and } a_k \text{ have the common endpoint } P_j. \end{cases}$$

(By our rather restricted definition of "graph" we have excluded the possibility of two edges with both endpoints in common.)

Finally it is easy to recognize that the groups of automorphisms of G and S are simply isomorphic, since any automorphism of G obviously induces one of S, and conversely.

That the lattice S is in general not distributive (nor even modular) may be shown by the following example. As graph G take that characterized by the quadratic form $x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1$, i.e., a quadrilateral with the edges

$$a_1 = P_1P_2,\ a_2 = P_2P_3,\ a_3 = P_3P_4,\ a_4 = P_1P_4.$$

The group of automorphisms of G is of course simply isomorphic to the octic group (= dihedral group of order 8). In the corresponding lattice S we have[5]

[4]In other words, S is the "cell-space" $P(G)$ of G (see *Lattice Theory*, 1st ed., p. 15) to which an O has been added in order to obtain a lattice.

[5]The "Hasse diagram" of this lattice may be obtained from the right-hand half of Fig 2, p. 15, of *Lattice Theory* by adding an O and joining it with B_1, B_2, B_3 and B_4.

$$I > A_i > B_i > O \qquad (i = 1, 2, 3, 4),$$

and also

$$A_1 > B_2,\ A_2 > B_3,\ A_3 > B_4,\ A_4 > B_1.$$

It is easily seen that this lattice S is not modular (hence not distributive). Indeed any modular lattice must satisfy the following condition (called (ξ') by Birkhoff[6]): "In a modular lattice, if X and Y cover[7] A, and $X \neq Y$, then $X \cup Y$ covers X and Y"; but the elements $X = B_1$ and $Y = B_3$ of S do not fulfil this condition.

We are now going to prove the following

THEOREM 2. *If $\mathfrak{G}$ is any abstract group of finite order g which can be generated by n of its elements, it is possible to find a lattice with at most $5(n + 2)\,g + 2$ elements whose group of automorphisms is simply isomorphic to $\mathfrak{G}$.*

Proof. It has been shown elsewhere[8] how to obtain a graph with at most $q = 2(n + 2)g$ vertices whose group of automorphisms is simply isomorphic to a given abstract group $\mathfrak{G}$; and since each of the vertices of that graph is of degree 3 (i.e., an endpoint of 3 edges), we have

$$p = 3q/2 = 3(n + 2)g.$$

With these values of p and q, Theorem 2 follows immediately from Theorem 1.

Of course it should be remarked that for special groups where a graph with fewer vertices and edges than the one used here is known, Theorem 1 will furnish a lattice with fewer elements than Theorem 2. E.g., for the octic group ($g = 8$, $n = 2$) Theorem 2 would give a lattice with 162 elements, but we know already that there is one with only 10 elements (see the example after the proof of Theorem 1).

Technical University Santa Maria,
Valparaiso (Chile)

---

[6]*Lattice Theory*, 1st ed., p. 34, Corollary 3 to Theorem 3.1.

[7]By "X covers A" it is meant that $X > A$, while no Z of S satisfies $X > Z > A$.

[8]R. Frucht, *Graphs of degree 3 with a given abstract group*. Can. J. Math., vol. I (1949), pp. 365-378.

A GENERALIZATION OF A THEOREM OF JACOBI ON SYSTEMS OF LINEAR DIFFERENTIAL EQUATIONS

CLYDE M. CRAMLET

JACOBI proved a curious theorem regarding the solutions of the system of equations

$$\frac{dx^1}{\lambda^1} = \frac{dx^2}{\lambda^2} = \ldots = \frac{dx^n}{\lambda^n},$$

for functions $\lambda^a(x^1, \ldots, x^n)$ satisfying

$$\frac{\partial \lambda^1}{\partial x^1} + \frac{\partial \lambda^2}{\partial x^2} + \ldots + \frac{\partial \lambda^n}{\partial x^n} = 0,$$

showing that the knowledge of $n-2$ independent integrals of the system leads, with this condition, to an exact differential equation for the last integral of the system. When the coordinates are Euclidean the left member is called the divergence of the vector λ^a. If the divergence of λ^a is non-vanishing there exists a factor M such that the divergence of $M\lambda_a$ vanishes. Jacobi's "theorem of the last multiplier"[1] states that the determination of this factor is tantamount to finding the last integral of the linear system.

Here a theorem is proved regarding a special system of k vectors, which we choose to call a Jacobian system of vectors. For $k = 1$ this theorem reduces to Jacobi's theorem of the last multiplier.

1. Conventions. The symbols $\lambda^a{}_{i|}(a = 1, \ldots, n;\ i = 1, \ldots, k)$ will represent functions of n independent variables $x = [x^1, \ldots, x^n]$. The ordered set of functions associated with a fixed i $(a = 1, \ldots, n)$ will be called a *vector*, k linearly independent vectors, a *basis*. A vector $a^r\lambda^a{}_{r|}$, the a's dependent on the x's, will be said to belong to the basis. The totality of vectors belonging to the basis constitutes a k-uple. Repeated Latin letters indicate a summation from 1 to k, repeated Greek, from 1 to n. All functions will be assumed to have such character as to satisfy the existence theorems that are applied. Only a finite number of derivatives need be assumed to exist in any case.

A coordinate transformation will be indicated formally by the equations

$$\bar{x}^a = \bar{x}^a(x) \qquad q \equiv \left|\frac{\partial \bar{x}}{\partial x}\right| \neq 0, \tag{1.1}$$

and the inverse by

Received August 18, 1949. Presented to the American Mathematical Society at the British Columbia meeting, June 19, 1948, under similar title.

[1]Goursat-Hedrick, *Mathematical Analysis*, Vol. II, Part II, Article 32.

$$x^\alpha = x^\alpha(\bar{x}) \qquad p = \left|\frac{\partial x}{\partial \bar{x}}\right|, \; pq = 1. \tag{1.2}$$

We shall have occasion to use the equations

$$\lambda^\alpha{}_{i|}\frac{\partial u}{\partial x^\alpha} = 0, \tag{1.3}$$

$$\frac{\partial x^1}{\lambda^1{}_{i|}} = \frac{\partial x^2}{\lambda^2{}_{i|}} = \ldots = \frac{dx^n}{\lambda^n{}_{i|}}\,, \tag{1.4}$$

$$\frac{\partial M\lambda^\alpha{}_{i|}}{\partial x^\alpha} = 0, \tag{1.5}$$

defining quantities u, and M under suitable conditions. When these exist they will be defined in a new coordinate system by the following conventions. The function $u(x)$ with the x's replaced from equations (1.2) will determine a function

$$\bar{u}(\bar{x}) = u(x), \tag{1.6}$$

which will represent the scalar u in the new coordinates x. The product $M(x)p$ with x replaced by (1.2) determines a representative $\overline{M}(\bar{x})$ in the new coordinate system

$$\overline{M}(\bar{x}) = M(x)p. \tag{1.7}$$

In this case M is said to be a *relative invariant of weight* 1.

Vectors $\bar{\lambda}^\alpha{}_{i|}$, representatives of $\lambda^\alpha{}_{i|}$, will be defined in a new coordinate system by the law of transformation of contravariant tensors: $\bar{\lambda}^\alpha{}_{i|} = \lambda^\rho{}_{i|}\dfrac{\partial \bar{x}^\alpha}{\partial x^\rho}$. With these conventions, the left members of (1.3) are invariant, and the left members of (1.5) are relative invariants of weight 1. The equations (1.3), (1.4) and (1.5) will imply like equations in new coordinates. If u and M are solutions of (1.3) and (1.5), $\bar{u}$ and $\overline{M}$ will be solutions of their representatives in the new coordinates. If $u = c$ is an integral of (1.4), $\bar{u} = c$ will be a representative integral in the new coordinates.

2. Complete basis. From the two contravariant vectors $\lambda^\alpha{}_{i|}$ and $\lambda^\alpha{}_{j|}$ an *associated contravariant* vector is defined by the equations

$$T^\beta{}_{ij|} \equiv \lambda^\alpha{}_{i|}\frac{\partial \lambda^\beta{}_{j|}}{\partial x^\alpha} - \lambda^\alpha{}_{j|}\frac{\partial \lambda^\beta{}_{i|}}{\partial x^\alpha}\,. \tag{2.1}$$

When the associate vectors of all pairs in a basis belong to the k-uple, the basis will be said to be *complete*. This is in agreement with the classical terminology that the system (1.3) is complete when the equations

$$T^\beta{}_{ij|}\frac{\partial u}{\partial x^\beta} = 0, \tag{2.2}$$

are dependent on (1.3). Similarly, when the associate vectors are null vectors the basis is said to be Jacobian. Some theorems in the theory of the linear systems of partial differential equations (1.3) will be restated in terms of these definitions:[3]

(2.3) *A system of k linearly independent vectors is always complete if $k = n$. If $k < n$ and the system is not complete, vectors $T^{\beta}{}_{ij|}$ may be adjoined to the system to form a set of $k' > k$ independent vectors. When the new system is not complete the process may be repeated until a complete system is obtained.* Completeness is a property of the k-uple.

(2.4) *A complete k-uple has bases that are Jacobian.* This is a property of a basis.

(2.5) *These properties are invariant under coordinate transformations.*

3. Normal form for a complete basis. The equation of (1.3) with $i = 1$ has $n - 1$ independent solutions $\phi^{A}(x)(A = 2, \ldots, n)$. Adjoin to these a function $\phi^{1}(x)$ such that the n functions are functionally independent. In new coordinates $\bar{x}^{\alpha} = \phi^{\alpha}(x)$ $(\alpha = 1, \ldots, n)$; this equation has solutions $\bar{x}^{A}$. Hence $\bar{\lambda}^{A}{}_{1|} = 0$. Since $\lambda^{\alpha}{}_{1|}$ is a non-null vector $\bar{\lambda}^{\alpha}{}_{1|}$ is non-null and $\bar{\lambda}^{1}{}_{1|} \neq 0$. Consequently there is no loss in generality in taking $\lambda^{1}{}_{1|}, 0, \ldots, 0$ as the components of the first vector in the original coordinates. By a subsequent transformation of coordinates $\bar{\lambda}^{\alpha}{}_{1|} = \lambda^{\beta}{}_{1|} \dfrac{\partial \bar{x}^{\alpha}}{\partial x^{\beta}} = \lambda^{1}{}_{1|} \dfrac{\partial \bar{x}^{\alpha}}{\partial x^{1}}$. By choosing $\bar{x}^{A}$ independent of x^{1} and $\bar{x}^{1} = \int dx^{1}/\lambda^{1}{}_{1|}$, the vector transforms to $1, 0, \ldots, 0$, and the corresponding equation takes the form $\dfrac{\partial u}{\partial x^{1}} = 0$.

Because of the hypothesis that the vectors form a complete basis the equations (1.3) have $n - k$ solutions ϕ^{A} $(A = k + 1, \ldots, n)$ that are now independent of x^{1}. Adjoining functions $\phi^{1} = x^{1}$, ϕ^{B} $(B = 2, \ldots, k)$ independent of x^{1}, so that $\phi^{\alpha}(\alpha = 1, \ldots, n)$ are independent, a transformation of coordinates may be defined by $\bar{x}^{\alpha} = \phi^{\alpha}(x)$. In the new coordinates the equations (1.3) are satisfied by $\bar{x}^{A}$, which implies that $\bar{\lambda}^{A}{}_{i|} = 0$. The components of the vector $1, 0, \ldots, 0$, are unchanged by this transformation. Hence:

(3.1) A *complete k-basis can be transformed to*

$$\lambda^{\alpha}{}_{1|} = \delta^{\alpha}{}_{1} \ (\alpha = 1, \ldots, n), \qquad \lambda^{\alpha}{}_{i|} = 0 \ (\alpha > k;\ i = 2, \ldots, k).$$

4. Normal form for a Jacobian system. It will be proved that:

(4.1) *A coordinate system exists in which a Jacobian system takes the normal form*

$$\lambda^{\alpha}{}_{i|} = \delta^{\alpha}{}_{i} \qquad (i = 1, \ldots, k;\ \alpha = 1, \ldots, n).$$

[3] Goursat-Hedrick, *op. cit.*, Section 89, p. 267.

To construct a proof by induction let $h - 1 < k$ of the vectors be assumed to be in the form of the theorem. The Jacobian condition $T^{\beta}{}_{ij|} = 0$, implies on some remaining vector $\lambda^{\alpha}{}_{h|}$ that:

$$\frac{\partial \lambda^{\alpha}{}_{h|}}{\partial x^{i}} = 0 \qquad (\alpha = 1, \ldots, n;\ i = 1, \ldots, h-1) \tag{4.2}$$

so that the components $\lambda^{\alpha}{}_{h|}$ are functions of $y \equiv [x^{h}, \ldots, x^{n}]$. The equations (1.4) $i = h$ have integrals $\phi^{\alpha} = c^{\alpha}$, $\alpha \neq h$ such that

$$\phi^{A} \equiv x^{A} - f^{A}(y) \qquad (A = 1, \ldots, h-1),$$

$$\phi^{A} \equiv \phi^{A}(y) \qquad (A = h+1, \ldots, h).$$

Let $\phi^{h}(y)$ be any function such that a proper transformation of coordinates may be defined by $\bar{x}^{\alpha} = \phi^{\alpha}(x, y)$. In the new coordinates only the hth component of $\lambda^{\alpha}{}_{h|}$ is non-vanishing, and it is a function of the variables y so may be reduced to unity by a transformation on these variables. These transformations do not affect the components of the vectors $\lambda^{\alpha}{}_{i|}(i = 1, \ldots, h-1)$. This completes the induction and the theorem follows.

5. Multipliers. A function M, satisfying an equation (1.5) has been called by Lagrange, a *multiplier* of the vector $\lambda^{\alpha}{}_{i|}$. In this case the vector $M\lambda^{\alpha}{}_{i|}$ is said to be solenoidal. To investigate the conditions that the system of k vectors admit the same multiplier, set $\mu_i \equiv -\dfrac{\partial \lambda^{\alpha}{}_{i|}}{\partial x^{\alpha}}$ and define the dependent variable M implicitly by an unknown function $Q(x, M) = 0$. The equations then take the homogeneous form

$$\lambda^{\alpha}{}_{i|} \frac{\partial Q}{\partial x^{\alpha}} + M\mu_i \frac{\partial Q}{\partial M} = 0. \tag{5.1}$$

Every solution Q of these equations that depends on M yields, with $Q = 0$, a solution M of (1.5). Every solution $M \equiv \phi(x)$ of (1.5) gives a $Q \equiv M - \phi(x)$ satisfying (5.1) for $Q = 0$. The problem of solving (1.5) for M therefore reduces to the problem of finding solutions of (5.1) that are dependent on M.

The completeness conditions of (5.1), the analogues of (2.2) are

$$T^{\beta}{}_{ij|} \frac{\partial Q}{\partial x^{\beta}} + Mt_{ij} \frac{\partial Q}{\partial M} = 0, \tag{5.2}$$

$$t_{ij} \equiv \lambda^{\alpha}{}_{i|} \frac{\partial \mu_j}{\partial x^{\alpha}} - \lambda^{\alpha}{}_{j|} \frac{\partial \mu_i}{\partial x^{\alpha}} \equiv -\frac{\partial}{\partial x^{\alpha}} T^{\alpha}{}_{ij|}.$$

The coefficients of $\dfrac{\partial Q}{\partial M}$ in (5.2) and in (5.1) are the same functions of the remaining differential coefficients, hence (5.2) may be assumed to be included in (5.1) which consequently, when integrable, may be assumed to be complete.

This requires that the basis $\lambda^a{}_{i|}$ is also complete. The converse is not true. But when the basis is complete and (5.1) is not complete an equation $\dfrac{\partial Q}{\partial M} = 0$ may be deduced as an essential condition on a solution of (5.1). These facts may be summarized in the theorem:

(5.3) *Sufficient conditions that a basis admit a multiplier are that the basis is complete, that is, that functions $a^r{}_{ij}$ exist such that*

$$T^a{}_{ij} \equiv a^r{}_{ij}\lambda^a{}_{r|},$$

and that these functions also reduce the equations

$$\frac{\partial T^\beta{}_{ij}}{\partial x^\beta} = a^r{}_{ij}\,\frac{\partial \lambda^\beta{}_{r|}}{\partial x^\beta},$$

to identities. These conditions are necessary when the basis has been completed.

These conditions are satisfied for Jacobian bases. The $a^r{}_{ij}$ being identically equal to zero, hence:

(5.4) *Each Jacobian basis of a complete k-uple admits a common multiple M such that the contravariant vectors of weight* 1, *$M\lambda^a{}_{i|}$ are solenoidal.*

6. Vector product. The vector product (non-metric) of $n - 1$ vectors may be defined by the covariant vector of weight -1 :

$$\lambda_a \equiv \epsilon_{aa_1 \ldots a_n}\lambda^{a_1}{}_{1|} \ldots \lambda^{a_k}{}_{k|}, \qquad k = n - 1. \tag{6.1}$$

For a scalar μ of weight 1, $\mu\lambda_a$ is a covariant vector and

$$a_{a\beta} \equiv \frac{\partial \mu\lambda_a}{\partial x^\beta} - \frac{\partial \mu\lambda_\beta}{\partial x^a}, \tag{6.2}$$

is the covariant tensor known as the curl. From (6.1) it appears that

$$\lambda^a{}_{i|}\lambda_a = 0 \qquad (i = 1, \ldots, k = \mu - 1). \tag{6.3}$$

Conversely these equations determine λ_a to within a factor of proportionality. By differentiating these the definition (6.2) leads to

$$\mu T^\beta{}_{ij|}\,\lambda_\beta \equiv a_{a\beta}\lambda^a{}_{i|}\,\lambda^\beta{}_{j|}\,. \tag{6.4}$$

The elimination of the factor μ from these equations gives

$$T^\beta{}_{ij|}\,\lambda_\beta \equiv \left(\frac{\partial \lambda_a}{\partial x^\beta} - \frac{\partial \lambda_\beta}{\partial x^a}\right)\lambda^a{}_{i|}\,\lambda^\beta{}_{j|}\,. \tag{6.5}$$

From (6.1) it is apparent that the vanishing of the left members of either of these sets of identities implies that the $(n - 1)$-uple be complete. For μ to be an integrating factor of $\lambda_a dx^a$ that is, for $\mu\lambda_a$ to be a gradient, it is necessary and sufficient that $a_{a\beta} = 0$. Then by (6.4) the basis is complete. For a com-

plete $n-1$ basis in normal form, by Theorem (3.1) $\lambda_a = 0$ ($a = 1, \ldots, n-1$), $\lambda_n \neq 0$. Choosing $\mu = \phi/\lambda_n$, ϕ an arbitrary function of x^n, $\mu\lambda_a$ is a gradient. Hence

(6.6) *The necessary and sufficient condition that the vector product of an $n-1$ basis be proportional to a gradient is that the basis be complete.*

This theorem may be stated in the equivalent form:

(6.7) *The necessary and sufficient conditions that the vector field λ_a be lamellar is that the basis be complete.*

The vector product of $k = n - 1$ gradients may be defined by the relative contravariant tensor of weight 1

$$\lambda^a = \epsilon^{a\sigma_1 \ldots \sigma_k} \frac{\partial u_1}{\partial x^{\sigma_1}} \cdots \frac{\partial u_n}{\partial x^{\sigma_k}} \tag{6.8}$$

where $u_1, \ldots, u_k$ are $n-1$ scalars. It is interesting to compare Theorem (6.7) with the well known theorem[3] that λ^a is solenoidal, and that any solenoidal vector is the vector product of $n-1$ gradients.

7. Generalization of a theorem of Jacobi. When the Jacobian system of $k = n - 1$ vectors $\lambda^a{}_{i|}$ is represented in the normal form (4.1), their vector product $\lambda_a \equiv \delta_{an}$ and all factors μ are given by $\mu \equiv \phi(x^n)$, ϕ being any integrable function. All multipliers of the basis are given by $M \equiv \phi(x^n)$ and therefore:

(7.1) *A Jacobian basis with $n-1$ vectors $\lambda^a{}_{i|}$ has multipliers M. For all such the vectors $M\lambda^a{}_{i|}$ are solenoidal and $M\lambda_a$ is a gradient. Conversely all factors M such that $M\lambda_a$ is a gradient imply that the vectors $M\lambda^a{}_{i|}$ are solenoidal.*

A system of contravariant vectors satisfying the hypotheses of (7.1) may be obtained as follows: Let $\phi^{k+2}, \ldots, \phi^n$ be $n-k-1$ integrals of a Jacobian system (1.3). Adjoin functions so that ϕ^a are n independent functions. The transformation $x^a = \phi^a(x)$ reduces this system to a Jacobian system of k equations in $k+1$ independent variables. With $k+1$ playing the role of n the conditions of the theorem are satisfied.

Let $\theta = C$ be the integral of the exact equation $\mu\lambda_a dx^a = 0$; then $\dfrac{\partial\theta}{\partial x^a} = \mu\lambda_a$. It follows from (6.3) that

$$\lambda^a{}_{i|} \frac{\partial\theta}{\partial x^a} \equiv 0, \tag{7.2}$$

and $\theta(x)$ is the "last" solution of the system (1.3). Although the index a is assumed to run from 1 to $k+1$ in these equations, it may as well run from

[3]Goursat-Hedrick, *op. cit.*

1 to n, the remaining terms vanishing. The equations are invariant and imply the following theorem:

(7.3) *Every system of equations of the form* (1.3) *is equivalent to a complete system*

$$\mu^{a}{}_{i|}(x)\frac{\partial u}{\partial x^{a}} = 0 \qquad (a = 1, \ldots, n, i = 1, \ldots, h < n),$$

such that the vectors $\mu^{a}{}_{i|}$ *admit a common multiplier* M *for which*

$$\frac{\partial}{\partial x^{a}}(M\mu^{a}{}_{i|}) = 0.$$

The system has $n - h$ *independent solutions: and a knowledge of* $n - h - 1$ *independent integrals, together with such a multiplier* M, *leads to an exact differential equation for the last solution.*

University of Washington,
Seattle

UNIFIED FIELD THEORY

MAX WYMAN

Introduction. In a recent unified theory originated by Einstein and Straus [1], the gravitational and electromagnetic fields are represented by a single non-symmetric tensor g_{ij} which is a function of four coordinates $x^r (r = 1, 2, 3, 4)$. In addition a non-symmetric linear connection $\Gamma_{jk}{}^i$ is assumed for the space and a Hamiltonian function is defined in terms of g_{ij} and $\Gamma_{jk}{}^i$. By means of a variational principle in which the g_{ij} and $\Gamma_{jk}{}^i$ are allowed to vary independently the field equations are obtained and can be written

$$g_{ik,a} - g_{sk}\Gamma_{ia}{}^s - g_{is}\Gamma_{ak}{}^s = 0, \tag{0.1}$$

$$\Gamma_{ia}{}^a - \Gamma_{ai}{}^a = 0, \tag{0.2}$$

$$R_{\underline{ik}} = 0, \tag{0.3}$$

$$R_{\underset{\vee}{ik},a} + R_{\underset{\vee}{ka},i} + R_{\underset{\vee}{ai},k} = 0. \tag{0.4}$$

In the above equations the comma in $g_{ik,a}$ or $R_{ik,a}$ denotes partial differentiation with respect to x^a. Further R_{ik} stands for the Ricci tensor based on the linear connection $\Gamma_{jk}{}^i$. The symbols $R_{\underline{ik}}$, $R_{\underset{\vee}{ik}}$ stand, respectively, for the symmetric and skew-symmetric parts of the tensor R_{ik} and hence

$$R_{\underline{ik}} = \tfrac{1}{2}(R_{ik} + R_{ki}), \tag{0.5}$$

$$R_{\underset{\vee}{ik}} = \tfrac{1}{2}(R_{ik} - R_{ki}). \tag{0.6}$$

The same notation is used throughout to denote the symmetric and skew-symmetric parts of other quantities entering into the new theory.

In the linearized field equations corresponding to the rigorous field equations (0.1)-(0.4) it has been found that the linearized field equations for the skew-symmetric part of the field are weaker than Maxwell's equations. It was pointed out that this in itself did not constitute a justified objection to the new theory as it was not known whether there were rigorous solutions of the field equations which were regular in all space and which would correspond to the solutions one could obtain for the linearized equations. For this reason it became important to determine rigorous solutions of equations (0.1)-(0.4).

Recently Papapetrou[1] has discussed the static spherically symmetric form of these equations and has discovered two rigorous solutions. The second

Received September 4, 1949.

[1]In [2] the field equations contain a cosmological constant λ which is zero in the Einstein field equations. When using Papapetrou's results we shall always take $\lambda = 0$.

solution is a very special case. In discussing his solutions Papapetrou points out that neither solution approaches asymptotically the corresponding solution obtained by means of the General Theory of Relativity.

In the present paper we shall generalize Papapetrou's second solution and shall in addition discuss some of the difficulties presented by the new Unified Theory.

1. Papapetrou's second case. Papapetrou took the static spherically symmetric tensor g_{ik} to have, in spherical polar coordinates, the form

$$g_{ik} = \begin{bmatrix} -\alpha & 0 & 0 & w \\ 0 & -\beta & r^2 v \sin\theta & 0 \\ 0 & -r^2 v \sin\theta & -\beta\sin^2\theta & 0 \\ -w & 0 & 0 & \gamma \end{bmatrix},$$

where α, β, γ, v, w are undetermined functions of r. For the case $v = 0$, $w \neq 0$, the general solution of the field equations was found to be [2]

$$\alpha = (1 - 2m/r)^{-1}, \; \beta = r^2,$$

$$\gamma = (1 + l^4/r^4)(1 - 2m/r), \; v = 0, \; w = \pm l^2/r^2,$$

where m, l are constants of integration. For the second case $v \neq 0$, $w = 0$ Papapetrou was unable to find the general solution but found a special case

$$\gamma = \alpha^{-1} = (1 - 2m/r), \; \beta = r^2, \; v = -c, \; w = 0,$$

where m, c are constants of integration. We shall now proceed to find the general solution corresponding to this second case $v \neq 0$, $w = 0$.

For the case $v \neq 0$, $w = 0$ Papapetrou has shown that the field equations reduce to

$$f = vr^2, \quad A = (\beta\beta' + ff')/(f^2 + \beta^2), \; B = (f\beta' - \beta f')/(f^2 + \beta^2), \tag{1.1}$$

$$A' + \tfrac{1}{2}(A^2 + B^2) - \tfrac{1}{2}A\,[(\alpha'/\alpha) + (\gamma'/\gamma)] = 0, \tag{1.2}$$

$$\gamma'' - \tfrac{1}{2}\gamma'[(\alpha'/\alpha) + (\gamma'/\gamma)] + A\gamma' = 0, \tag{1.3}$$

$$\beta'' - f'B - \tfrac{1}{2}\beta'[(\alpha'/\alpha) - (\gamma'/\gamma)] + 2\alpha(2\beta fc - \beta^2 + f^2)/(f^2 + \beta^2) = 0, \tag{1.4}$$

$$f'' + \beta'B - \tfrac{1}{2}f'[(\alpha'/\alpha) - (\gamma'/\gamma)] - 2\alpha(2\beta f + c\beta^2 - cf^2)/(f^2 + \beta^2) = 0, \tag{1.5}$$

where the prime notation indicates differentiation with respect to r and c is an arbitrary constant of integration. In the above equations (1.1) is simply a definition of the symbols A, B and f while the remaining equations are the field equations for this particular case.

Since $A = \dfrac{d}{dr}\log(f^2 + \beta^2)^{\frac{1}{2}}$ equation (1.3) can be integrated to give

$$\gamma' = 2m\,[\alpha\gamma/(f^2 + \beta^2)]^{\frac{1}{2}}, \tag{1.6}$$

where m is an arbitrary constant of integration. It has been taken in this

form as it will later be identified with the mass of the spherical body. We shall throughout the remainder of this section deal only with the case $m \neq 0$. When this is so $\gamma' \neq 0$ and γ is not a constant.

Due to the tensorial character of g_{ik} one of α, β, γ can be chosen arbitrarily.[3] We shall find that the general solution is most easily obtained if we allow γ to be the variable that has this arbitrary character.

Concentrating our attention on equations (1.4) and (1.5) we find it advantageous to replace these equations by two equivalent equations. Multiplying (1.4) by $\beta/(f^2+\beta^2)$ and (1.5) by $f/(f^2+\beta^2)$ and adding the results we find

$$(1.7)\quad (\beta\beta''+ff'')/(f^2+\beta^2)+B^2-\tfrac{1}{2}A[(\alpha'/\alpha)-(\gamma'/\gamma)]+2\alpha(cf-\beta)/(f^2+\beta^2)=0.$$

Since

$$A'=(\beta\beta''+ff'')/(f^2+\beta^2)+B^2-A^2,$$

(1.7) can be written

$$(1.8)\quad A'+A^2-\tfrac{1}{2}A[(\alpha'/\alpha)-(\gamma'/\gamma)]+2\alpha(cf-\beta)/(f^2+\beta^2)=0.$$

Similarly by multiplying (1.4) by $f/(f^2+\beta^2)$ and (1.5) by $\beta/(f^2+\beta^2)$ and subtracting the results we can obtain the equation

$$(1.9)\quad B'+AB-\tfrac{1}{2}B[(\alpha'/\alpha)-(\gamma'/\gamma)]+2\alpha(c\beta+f)/(f^2+\beta^2)=0.$$

Thus equations (1.8) and (1.9) are equivalent to (1.4) and (1.5).

If we let $i=(-1)^{\frac{1}{2}}$ and introduce the complex variable $q=k+iu$ by means of the equation

$$(1.10)\qquad f+i\beta=e^q,$$

we find that

$$(1.11)\qquad A+iB=q',$$

and hence $A=k'$, $B=u'$. Multiplying (1.9) by i and adding (1.8) one obtains the equation

$$(1.12)\quad q''+[A-\tfrac{1}{2}\{(\alpha'/\alpha)-(\gamma'/\gamma)\}]q'+2\alpha(c+i)e^q/(f^2+\beta^2)=0.$$

Thus the single equation (1.12) in the complex variable q is equivalent to the two real equations (1.8) and (1.9).

Since m was assumed to be non-zero we can solve (1.6) for α to obtain

$$(1.13)\qquad \alpha=(\gamma')^2(f^2+\beta^2)/4m^2\gamma.$$

Substituting in (1.12) for α gives

$$(1.14)\quad q''-q'[(\gamma''/\gamma')-(\gamma'/\gamma)]+\gamma'^2(c+i)e^q/2m^2\gamma=0.$$

[3]Since we are excluding the case γ = constant our phrase "chosen arbitrarily" excludes this choice of γ for which the statement is not true. Certain differentiability conditions are also implied by the field equations.

From the fact that $q' = \frac{dq}{d\gamma}\gamma'$ and $q'' = \frac{d^2q}{d\gamma^2}\gamma'^2 + \frac{dq}{d\gamma}\gamma''$, equation (1.14) can be written

$$\frac{d^2q}{d\gamma^2} + \left(\frac{dq}{d\gamma}/\gamma\right) + (c+i)e^q/2m^2\gamma = 0. \tag{1.15}$$

The substitution

$$q = y - \log\gamma, \quad x = \log\gamma, \tag{1.16}$$

reduces this equation to

$$\frac{d^2y}{dx^2} + [(c+i)e^y/2m^2] = 0. \tag{1.17}$$

Equation (1.17) is easily integrated once to give

$$\left(\frac{dy}{dx}\right)^2 + [(c+i)e^y/m^2] = h, \tag{1.18}$$

where h is an arbitrary complex constant of integration. In (1.18) we can separate the variables and then integrate to find

$$e^y = [m^2h \operatorname{sech}^2(\tfrac{1}{2}h^{\frac{1}{2}}x + a)]/(c+i), \tag{1.19}$$

where a is a second complex constant of integration. Returning to our original variable we find

$$e^q = 4m^2h/[(e^a\gamma^{\frac{1}{2}h^{\frac{1}{2}}} + e^{-a}\gamma^{-\frac{1}{2}h^{\frac{1}{2}}})^2\gamma(c+i)]. \tag{1.20}$$

Thus far we have found the general solution of equations (1.3), (1.4) and (1.5) and so far no use has been made of equation (1.2). From the tensorial character of our equations and the arbitrary character of γ we know that one of the equations (1.2), (1.3), (1.4) and (1.5) is redundant. It has however been shown by Papapetrou that this redundant equation is (1.5). We shall see that in order for our solution (1.20) to satisfy (1.2) the number of arbitrary constants in the solution is reduced by one.

If we transform equation (1.18) back to the variables q and γ by means of $y = q + \log\gamma$, $x = \log\gamma$ we find

$$\left(\gamma\frac{dq}{d\gamma} + 1\right)^2 + [(c+i)e^q\gamma/m^2] = h. \tag{1.21}$$

Since $\frac{dq}{d\gamma} = q'/\gamma'$ this equation can be written

$$(q')^2 + (2\gamma'q'/\gamma) + [(c+i)e^q\gamma'^2/m^2\gamma] = (h-1)\gamma'^2/\gamma^2. \tag{1.22}$$

Substituting $q' = A + iB$ and $e^q = f + i\beta$ we can by equating real and imaginary parts of (1.22) obtain the equations

$$A^2 - B^2 + [2\gamma' A/\gamma] + (cf - \beta)\gamma'^2/m^2\gamma = (h_0 - 1)\gamma'^2/\gamma^2, \tag{1.23}$$

$$AB + (\gamma' B/\gamma) + [(c\beta + f)\gamma'^2/2m^2\gamma] = h_1\gamma'^2/2\gamma^2, \tag{1.24}$$

where the complex constant h has been written $h = h_0 + ih_1$. If we multiply (1.23) by $\frac{1}{2}$ and subtract the result from (1.8) we have

$$A' + \tfrac{1}{2}(A^2 + B^2) - \tfrac{1}{2}A[(\alpha'/\alpha) + (\gamma'/\gamma)] = (1 - h_0)\gamma'^2/2\gamma^2. \tag{1.25}$$

Hence (1.2) will be satisfied only if $h_0 = 1$.

Thus for the case $m \neq 0$ the general solution of the field equations is

$$f + i\beta = 4m^2h/[(e^{a}\gamma^{\frac{1}{2}h^{\frac{1}{2}}} + e^{-a}\gamma^{-\frac{1}{2}h^{\frac{1}{2}}})^2\gamma(c + i)] \tag{1.26}$$

$$\alpha = \gamma'^2(f^2 + \beta^2)/4m^2\gamma \tag{1.27}$$

where γ can be chosen to be any arbitrary function of r that we please and f, β are obtained by equating real and imaginary parts of (1.26). It is well to note that m, c are real arbitrary constants, that h has the form $h = 1 + ih_1$, and a is an arbitrary complex constant of integration.

Finally it is of interest to see that Papapetrou's special solutions result from the choice $\gamma = 1 - 2m/r$, $h_1 = 0$ and $e^{2a} = -1$.

We should at this point go on to see how the boundary conditions at infinity allow us to evaluate the arbitrary constants of our solution. However since there is a difficulty in choosing suitable boundary conditions, which we would like to present in some detail, we shall postpone this discussion to a later section.

2. Case $m = 0$. When the constant m is taken to be zero, equation (1.6) becomes $\gamma' = 0$. Thus γ is a constant and can be taken equal to one without loss of generality. Equation (1.12) is still valid and hence for $\gamma' = 0$ becomes

$$q'' + (A - \tfrac{1}{2}\alpha'/\alpha)q' + 2\alpha(c + i)e^q/(f^2 + \beta^2) = 0. \tag{2.1}$$

Multiplying by $(f^2 + \beta^2)q'/\alpha$ we can immediately integrate once with respect to r to give

$$q'^2 + 4\alpha(c + i)e^q/(f^2 + \beta^2) = 4h\alpha/(f^2 + \beta^2), \tag{2.2}$$

where h is an arbitrary complex constant of integration. From the tensorial character of g_{ik} we know that we can make any transformation of the form $r = r(x)$ without destroying the relationship $\gamma = 1$. Thus if we make the transformation

$$x = \int [\alpha/(f^2 + \beta^2)]^{\frac{1}{2}}\, dr \tag{2.3}$$

then $\alpha/(f^2 + \beta^2) = \left(\dfrac{dx}{dr}\right)^2$ and (2.2) can be written

$$\left(\frac{dq}{dx}\right)^2 + 4(c + i)e^q = 4h. \tag{2.4}$$

The solution of this equation is

$$e^q = [h \operatorname{sech}^2(h^{\frac{1}{2}}x + a)]/(c + i) \tag{2.5}$$

if $h \neq 0$, and is

$$e^q = (i - c)/[(c^2 + 1)(x + a)^2] \tag{2.6}$$

if $h = 0$. In either case a is an arbitrary complex constant of integration. At this stage we have not ensured that equation (1.2) is satisfied. By an analysis similar to that used in the previous section we find that this will be so only if the constant h has the form $h = ih_0$. Thus the case $m = 0$ leads to the two possibilities

$$\left\{\begin{aligned} f + i\beta &= [h \operatorname{sech}^2(h^{\frac{1}{2}}x + a)]/(c + i), \\ \gamma &= 1, \\ \alpha &= (f^2 + \beta^2)\left(\frac{dx}{dr}\right)^2, \end{aligned}\right. \tag{2.7}$$

$$\left\{\begin{aligned} f + i\beta &= (i - c)/[(c^2 + 1)(x + a)^2], \\ \gamma &= 1, \\ \alpha &= (f^2 + \beta^2)\left(\frac{dx}{dr}\right)^2, \end{aligned}\right. \tag{2.8}$$

where in each case x can be any arbitrary function of r.

We shall again leave the discussion of the implications of the boundary conditions to a later section.

3. The metric of space-time. In the General Theory of Relativity we assume at the outset a four dimensional Riemannian space which of course implies the existence of a metric tensor which determines the properties of space-time. When the equations of motion of a particle are considered, the derivatives of the metric tensor a_{ik} enter in such a way that the components a_{ik} appear as gravitational potentials. This dual character of the metric tensor arises quite naturally and leads to no ambiguity. In the new theory the point of view has been altered. We assume at the outset certain field quantities g_{ij}, $\Gamma_{jk}{}^i$ and then derive field equations which will determine these field quantities. If we interpret the tensor g_{ij} as a representation of the combined gravitational and electromagnetic fields the question arises as to how the results of the new theory compare with the corresponding results of the General Theory of Relativity. Before this question can be answered we must in some way introduce a metric for space-time so that corresponding results can be compared.

It is natural to assume that at any point in space the symmetric metric tensor a_{ij} will be completely determined by our field quantities. This implies that the components a_{ij} will be certain functions of g_{ij}, $\Gamma_{jk}{}^i$. We denote this functional relationship by

$$a_{ij} = f_{ij}(g_{rs}, \Gamma_{rs}{}^p). \tag{3.1}$$

The field equations determine the quantities $\Gamma_{rs}{}^{p}$ in terms of g_{rs} and their first derivatives. Thus the above assumption is equivalent to saying that the metric tensor becomes completely determined at any point in space by a knowledge of the g_{rs} and their first derivatives.

The functional relationship of (3.1) is not quite arbitrary in that the components a_{ik} must be the components of a tensor. It is not too difficult to show that the allowable functions f_{ij} must satisfy certain partial differential equations in order for this to be true. Since however the field quantities g_{rs}, $\Gamma_{rs}{}^{p}$ determine the tensors $g_{\underline{rs}}$, $g_{\underset{\vee}{rs}}$, $g^{\underline{rs}}$, $g^{\underset{\vee}{rs}}$, $\Gamma_{\underset{\vee}{rs}}{}^{p}$ we can construct an infinity of tensors of the form (3.1).

The field equations of the Unified Theory reduce to those of General Relativity if $g_{\underset{\vee}{ij}} = 0$. Hence we shall make the requirement that

$$f_{ij}(g_{\underline{rs}}, \Gamma_{\underline{rs}}) = g_{\underline{rs}}, \quad \text{if } g_{\underset{\vee}{ij}} = 0.$$

At this stage of the theory there seems to be little to guide us in a suitable choice of metric tensor. However when one considers the equations of motion a strong argument can be advanced for a particular choice for the metric tensor.

Since the linear connection $\Gamma_{jk}{}^{i}$ has been assumed to be the linear connection by means of which we define the parallel displacement of a vector it seems natural to require that the equations of motion of a free particle can be put into the form

$$\frac{d^2x^i}{ds^2} + \Gamma_{jk}{}^{i}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0, \tag{3.2}$$

where s is a suitable parameter along the trajectory of motion. Because of the symmetry of $\dfrac{dx^j}{ds}\dfrac{dx^k}{ds}$ in the indices j, k the skew-symmetric part of the second term will cancel out and the equations of motion have the form

$$\frac{d^2x^i}{ds^2} + \Gamma_{\underline{jk}}{}^{i}\frac{dx^j}{ds}\frac{dx}{ds} = 0. \tag{3.3}$$

Multiplying (3.3) by $g_{\underline{im}}\dfrac{dx^m}{ds}$ and summing with respect to i we obtain

$$g_{\underline{im}}\frac{dx^m}{ds}\frac{d^2x^i}{ds^2} + g_{\underline{im}}\Gamma_{\underline{jk}}{}^{i}\frac{dx^j}{ds}\frac{dx^k}{ds}\frac{dx^m}{ds} = 0. \tag{3.4}$$

This can be put into the form

$$\frac{d}{ds}\left(g_{\underline{im}}\frac{dx^m}{ds}\frac{dx^i}{ds}\right) - g_{\underline{jk}/m}\frac{dx^j}{ds}\frac{dx^k}{ds}\frac{dx^m}{ds} = 0, \tag{3.5}$$

where $g_{jk/m}$ means the covariant derivative of $g_{\underline{jk}}$ with respect to the symmetric linear connection $\Gamma_{\underline{jk}}{}^{i}$. Thus

$$g_{\underline{im}}\frac{dx^i}{ds}\frac{dx^m}{ds} = \text{constant} \tag{3.6}$$

will be an integral of (3.5) providing we can show that

$$g_{\underline{jk}/m} + g_{\underline{km}/j} + g_{\underline{mj}/k} = 0. \tag{3.7}$$

These relations, we shall show, are an immediate consequence of equations (1) given in the introduction of our paper. From equations (1) we have

$$\begin{aligned} g_{\underline{ik},a} &= \tfrac{1}{2}(g_{sk}\Gamma_{ia}{}^{s} + g_{is}\Gamma_{ak}{}^{s} + g_{si}\Gamma_{ka}{}^{s} + g_{ks}\Gamma_{ai}{}^{s}) \\ &= g_{\underline{sk}}\Gamma_{\underline{ia}}{}^{s} + g_{\underset{\vee}{sk}}\Gamma_{\underset{\vee}{ia}}{}^{s} + g_{\underline{is}}\Gamma_{\underline{ak}}{}^{s} + g_{\underset{\vee}{is}}\Gamma_{\underset{\vee}{ak}}{}^{s}, \end{aligned}$$

and hence

$$g_{\underline{ik}/a} = g_{\underset{\vee}{sk}}\Gamma_{\underset{\vee}{ia}}{}^{s} + g_{\underset{\vee}{is}}\Gamma_{\underset{\vee}{ak}}{}^{s}. \tag{3.8}$$

Applying two cyclic permutations to the indices i, k, a, and adding the results to (3.8) we immediately find

$$g_{\underline{ik}/a} + g_{\underline{ka}/i} + g_{\underline{ai}/k} = 0. \tag{3.9}$$

Equations (3.9) are of course equivalent to (3.7).

Since the equations of motion (3.2) always have the quadratic expression (3.6) as an integral it seems natural to assume that the metric of space time is given by $g_{\underline{ij}}\, dx^i\, dx^j$ and hence our choice of metric tensor would have to be $a_{ij} = g_{\underline{ij}}$ even though $g_{\underset{\vee}{ij}} \neq 0$. Papapetrou has used the requirement $a_{ij} = g_{\underline{ij}}$ in connection with his solution of the field equations and has found the results of the new theory do not agree with those given by the General Theory of Relativity. He attributes this difficulty to the uncertainty in the physical identification of the tensor $g_{\underset{\vee}{ij}}$. While this may be true we feel that other possibilities exist. For example it might be that equations (3.2) are not the true equations of motion and that other equations will replace them. In this case it might be that the requirement $a_{ij} = g_{\underline{ij}}$ is an approximation and that the true metric will involve all our fundamental field quantities. In order to show the possibilities that exist we will examine the physical consequences when a different choice of metric is made.

Let us define two covariant vectors h_i, u_i by means of

$$h_i = g_{\underline{ab}}\, g^{\underset{\vee}{ba}}\Gamma_{\underset{\vee}{ai}}{}^{b}, \tag{3.10}$$

$$u_i = h_i/(g^{rs}h_r h_s)^{\frac{1}{2}} = h_i/(g^{\underline{rs}}h_r h_s)^{\frac{1}{2}}. \tag{3.11}$$

In (3.3) if h_i turns out to be a zero vector (i.e. $h_i = 0$) we simply take $u_i = 0$. There is of course a possibility that $g^{rs}h_r h_s \equiv 0$ with $h_i \neq 0$. We shall discuss this possibility at the end of this section. Finally we note that $g^{rs}h_r h_s$ could be negative and hence u_i would become an imaginary tensor. However (3.3) is only an intermediate step in our calculations and we shall see that this difficulty is removed with our final choice of metric.

We define a third covariant vector q_i by means of

$$q_i = (g_{im}\, g^{\underline{mn}}\, u_n)/[1 + \tfrac{1}{2} g_{\underset{\vee}{rs}}\, g^{\underset{\vee}{sr}}]^{\frac{1}{2}},$$

and then choose the metric

$$a_{ij} = g_{\underline{ij}} + q_i q_j. \tag{3.12}$$

Referring back to Papapetrou's exact solution as given in §1 we have that the non-zero components of g_{ij} are

$$g_{11} = -[1 - (2m/r)]^{-1},\ g_{22} = -r^2,\ g_{33} = -r^2 \sin^2 \theta,$$

$$g_{44} = [1 + (l^4/r^4)][1 - (2m/r)],\ g_{14} = -g_{41} = \pm l^2/r^2.$$

From these we can calculate the non-zero components of g^{ij} to be $g^{11} = -g_{44}$, $g^{22} = -1/r^2$, $g^{33} = -1/r^2 \sin^2 \theta$, $g^{44} = -g_{11}$, $g^{14} = -g^{41} = -g_{14}$. The non-vanishing components of $\Gamma_{\underset{\vee}{jk}}{}^i$ are

$$\Gamma_{\underset{\vee}{42}}{}^2 = -\Gamma_{\underset{\vee}{24}}{}^2 = -w/rg_{11} = \Gamma_{\underset{\vee}{43}}{}^3 = -\Gamma_{\underset{\vee}{34}}{}^3,\ \Gamma_{\underset{\vee}{14}}{}^1 = -\Gamma_{\underset{\vee}{41}}{}^1 = -2w/rg_{11}.$$

From these we can compute the components of h_i to be $[-l^4/r^5, 0, 0, 0]$. Hence the components of u_i are $[(g^{11})^{-\frac{1}{2}}, 0, 0, 0]$ and of q_i are $[0, 0, 0, g_{41}(g^{11})^{\frac{1}{2}}/(1 + g_{14}^2)^{\frac{1}{2}}]$. This finally gives the metric

$$a_{ij} = g_{\underline{ij}}, \quad \text{if } i, j \text{ are not both equal to 4},$$

$$a_{44} = g_{44} + g_{41}^2 g^{11}/(1 + g_{14}^2).$$

Substituting the values of the g's we find that the non-vanishing components of the metric tensor a_{ij} are given by

$$a_{11} = -(1 - 2m/r)^{-1}, \quad a_{22} = -r^2, \quad a_{33} = -r^2 \sin^2 \theta, \quad a_{44} = 1 - 2m/r.$$

This is of course the Schwarzschild solution of General Relativity.

We are not advocating the choice of metric (3.4) because it has been constructed in a very artificial manner. We use it to illustrate the importance of the choice of metric and to discuss several important points. If we assume that the metric of (3.4) is the true metric then we have seen the line element corresponding to Papapetrou's solution of the field equations is the Schwarzschild line element for a spherical mass with zero charge. Thus under this particular choice of metric we would have to say that Papapetrou's solution of the field equations is still a solution which corresponds to a pure gravitational field even though a second constant of integration l appears in the solution. This constant completely disappears when the components of the metric tensor are evaluated. Since $g_{\underset{\vee}{ij}} \neq 0$ in Papapetrou's solution our choice of metric also implies that $g_{\underset{\vee}{ij}}$ cannot be interpreted in terms of the electromagnetic field alone or else there exist electromagnetic fields which do not influence our measurements of space-time. This latter conclusion seems hardly likely and hence our example would seem to strengthen Papapetrou's conclusion that the physical interpretation of $g_{\underset{\vee}{ij}}$ is an open question. Finally we might point out that the disappearance of a constant of integration by

choice of metric may be connected with the fact that the linearized equations of the Unified Theory are weaker than Maxwell's equations. It might be possible that a choice of metric exists which make these weaker equations equivalent to Maxwell's equations.

If we agree that Papapetrou's choice of metric $a_{ij} = g_{\underline{ij}}$ is at best an approximation to the true metric then of course the accuracy of this approximation must be discussed. It is not difficult to construct metrics in which this approximation is valid only up to and including terms of the order $1/r$. Since it is the terms of order $1/r^2$ which measure the electromagnetic effect on space-time we see, for such metrics, that Papapetrou's approximation is equivalent to assuming a zero electromagnetic field. This then would be the reason that Papapetrou's solution does not behave asymptotically in the same manner as the solution in General Relativity corresponding to a point charge in which the terms of order $1/r^2$ are retained. It is very easy to construct metrics which show the same asymptotic behaviour as the General Relativity solutions up to and including terms of order $1/r^2$. However our construction is still very artificial and we shall not include this work in this paper.

In our derivation of the metric (3.4) we left in abeyance the possibility that $g^{rs}h_r h_s = 0$ with $h_r \neq 0$. For the static case, in which there exists a coordinate system in which the g_{rs} are all independent of the time-like coordinate x^4, it is possible to show, under suitable restrictions, that $g^{rs}h_r h_s = 0$ implies $h_r = 0$. We have not studied the non-static case in detail because we doubt very much that (3.4) will provide a suitable choice of metric. This section has been used only to show that a problem exists in the choice of a metric and that some logical physical reason should be advanced for the choice of metric for our new theory.

To conclude this section of the paper we would like to anticipate one criticism that might be made. It might be argued that the analogy from General Relativity would allow us to assume the dual nature of the tensor g_{ij}. By this I mean that the metric in space should be determined as a function of this tensor alone and would be independent of Γ_{jk}^{i}. Although this may be true it still does not destroy the point that we have been trying to make in this section. Out of such a tensor an unlimited number of metric tensors can be constructed and we must still advance some reason for a particular choice. As an example we might choose $a_{ij} = g^{\underline{mn}} g_{im} g_{jn}$. This particular metric turns out to be completely equivalent to Papapetrou's metric $a_{ij} = g_{\underline{ij}}$ for Papapetrou's particular spherically symmetric solution.

4. The boundary conditions. Since our field equations reduce to those of General Relativity if $g_{\check{ij}} = 0$ it is natural to assume that when $g_{\check{ij}} = 0$ our field is purely gravitational. Thus as boundary conditions it is natural to assume, in the general case, that at large distances from matter or charge there will exist a coordinate system in which the components of the metric tensor approach the scheme given by

$$(4.1)\qquad g_{ij} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We shall denote the scheme of (4.1) by γ_{ij} and as is usual we shall call this the Galilean tensor. In any other coordinate system the components of γ_{ij} are of course obtained by the tensor law of transformation. We notice of course that γ_{ij} provides an exact solution of our field equations in which $\Gamma_{jk}{}^{i} = 0$ in the coordinate system used for (4.1). This tensor is taken as the mathematical representation of the absence of both gravitational and electromagnetic fields. If we use the transformation

$$x^1 = r \sin\theta \cos\phi,\ x^2 = r\sin\theta\sin\phi,\ x^3 = r\cos\theta,\ x^4 = x^4,$$

the components $\bar{\gamma}_{ij}$ of the Galilean tensor are given by

$$(4.2)\qquad \bar{\gamma}_{ij} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -r^2 & 0 & 0 \\ 0 & 0 & -r^2\sin^2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

It would be in keeping with the principle of relativity if the condition that $g_{ij} \to \gamma_{ij}$ in one coordinate system implied that this was true in every coordinate system. Unfortunately this is not true and in fact it was used as a criticism of General Relativity when that theory was first proposed. For the General Theory of Relativity this difficulty was, in a sense, resolved for spherically symmetric solutions of the field equations, by means of Birkhoff's Theorem. Since the approach of a tensor to its Galilean values is not an invariant condition we must then single out a particular coordinate system if this condition is to be used as a boundary condition. We shall show by using our second solution of the field equations that this singling out of a special coordinate system presents a real difficulty in our new theory.

Papapetrou [2, p. 70] has shown that the general spherically symmetric form of $\underset{\vee}{g}_{ij}$ in Cartesian coordinates is

$$(4.3)\qquad \underset{\vee}{g}_{ij} = \begin{bmatrix} 0 & \frac{z}{r}v & -\frac{y}{r}v & \frac{x}{r}w \\ -\frac{z}{r}v & 0 & \frac{x}{r}v & \frac{y}{r}w \\ \frac{y}{r}v & -\frac{x}{r}v & 0 & \frac{z}{r}w \\ -\frac{x}{r}w & -\frac{y}{r}w & -\frac{z}{r}w & 0 \end{bmatrix},$$

where v, w are functions of r alone. Hence $\underset{\vee}{g}_{ij} \to 0$ as $r \to \infty$ implies that $v \to 0$ and $w \to 0$ as $r \to \infty$. In spherical polar coordinates the components $\underset{\vee}{g}_{ij}$ of this tensor are given by

$$(4.4) \qquad \bar{g}_{\underset{\vee}{ij}} = \begin{bmatrix} 0 & 0 & 0 & w \\ 0 & 0 & r^2 v \sin\theta & 0 \\ 0 & -r^2 v \sin\theta & 0 & 0 \\ -w & 0 & 0 & 0 \end{bmatrix}.$$

Hence the same conditions in this coordinate system imply $r^2 v \to 0$, $w \to 0$ as $r \to \infty$. This of course is a much stronger condition than the corresponding condition in Cartesian coordinates. We shall now use our second solution to show that these conditions imply different solutions of the field equation.

Returning to the solution given by (1.26), (1.27) our complete boundary conditions are

$$\alpha \to 1,\ \beta \to r^2,\ \gamma \to 1,\ f = vr^2 \to 0 \text{ as } r \to \infty$$

or

$$\alpha \to 1,\ \beta \to r^2,\ \gamma \to 1,\ v \to 0 \qquad \text{as } r \to \infty,$$

depending on the coordinate system used. Since $\beta \to \infty$ as $\gamma \to 1$ we must have from (1.26) that $e^a + e^{-a} = 0$. Thus $e^{2a} = -1$ and (1.26) can be written

$$(4.5) \qquad f + i\beta = -4m^2 h\, \gamma^{h^{\frac{1}{2}}-1}/(\gamma^{h^{\frac{1}{2}}} - 1)^2\,(c + i).$$

Moreover if we let $\gamma = 1 - x$ and expand (4.5) in terms of x we find

$$f + i\beta = \frac{4m^2(i - c)}{(c^2 + 1)x^2}\left[1 - (h - 1)\frac{x^2}{12} + \ldots\right].$$

Remembering that $h = 1 + ih_1$, we can equate real and imaginary parts to find

$$\beta = \frac{4m^2}{(c^2 + 1)x^2} + \frac{m^2}{3(c^2 + 1)}\, ch_1 + O(x),$$

$$f = -\frac{4m^2 c}{(c^2 + 1)x^2} + \frac{h_1 m^2}{3(c^2 + 1)} + O(x),$$

where we mean by $O(x)$ terms of the order x. As $\gamma \to 1$, $x \to 0$. Hence $f \to 0$ as $x \to 0$ only if $c = 0$ and $h_1 = 0$. In this solution $m = 0$ is not possible. Since $h_1 = 0$ implies $h = 1$ we find (4.5) becomes

$$(4.6) \qquad f + i\beta = +4m^2 i/(\gamma - 1)^2.$$

From the fact that the right side is a pure imaginary we can conclude that the strong boundary conditions result in $f \equiv 0$ and hence $g_{\underset{\vee}{ij}} \equiv 0$ and our resulting solution degenerates into the Schwarzschild solution.

If we use the weaker boundary condition that $v = f/r^2 \to 0$ as $r \to \infty$, we still find that $c = 0$ but we no longer have the condition that $h_1 = 0$. Thus in our final solution two arbitrary constants m, h_1 remain which can be interpreted as being determined by the mass and charge of the particle. Thus we see that the requirement that the tensor g_{ij} approach its Galilean values as

$r \to \infty$ implies different solutions in the two coordinate systems. As a matter of preference I feel the stronger boundary conditions will prove correct and that the solution we have obtained degenerates into the Schwarzschild solution for a pure gravitational field. I feel that the physical problem of a charged particle will only be solved when the general field equations are solved under the more general conditions $vw \neq 0$. The main reason for this belief is that we have shown that the solution resulting from the assumption $v = 0$, $w \neq 0$ can be interpreted under proper choice of metric, as being equivalent to the assumption $v = w = 0$. Similarly, under the strong boundary conditions, we have shown that the solution resulting from the assumption $w = 0$, $v \neq 0$ also degenerates to the case $v = w = 0$. For this reason it is possible that either of the restrictions imposed by Papapetrou, namely $v = 0$, $w \neq 0$, or $v \neq 0$, $w = 0$, may be equivalent to destroying the electromagnetic field.

For our solution of the field equations corresponding to the case $m = 0$ we can by similar analysis to that used in the present section show that the strong boundary conditions reduce this solution to that for zero mass and zero charge.

5. Conclusion. At the present stage our theory is still far from complete. A proper choice of metric has not been made nor have the equations of motion of a particle been defined. It seems necessary, therefore, to study the physical significance of our field quantities so that the present theory can be completed in a logical manner. When this is done it seems likely that the difficulties raised in the present paper will be removed.

References

[1] A. Einstein and E. G. Straus, *A generalisation of the Relativistic Theory of Gravitation* II, Ann. of Math., vol. 47 (1946), 731-741.

[2] A. Papapetrou, *Static spherically symmetric solutions in the unitary field theory*, Proc. Royal Irish Acad., vol. A52 (1948), 69-96.

University of Alberta

EQUATION DE HILL ET PROBLEME DE STÖRMER

RENE DE VOGELAERE

1. Introduction. Pour la détermination des orbites infiniment voisines de l'équateur, dans le problème de Störmer, une équation de Hill est à résoudre. Les méthodes sont expliquées d'abord sur l'équation générale, puis appliquées au problème de Störmer. Signalons les résultats suivants: les orbites équatoriales, considérées du point de vue de leur perturbation dans le plan méridien sont successivement stables, instables impaires, stables, instables paires et cela indéfiniment quand γ_1 se rapproche de un; quelques orbites limites entre les zones de stabilité et d'instabilité sont obtenues avec une méthode qui permet n'importe quelle précision désirée.

EQUATION DE HILL

1. Généralités. C'est une équation du type

$$\frac{d^2\eta}{d\sigma^2} + f(\sigma)\eta = 0 \tag{1.1}$$

où $f(\sigma)$ est une fonction périodique de σ, [5] et [17]. La solution peut se mettre sous la forme

$$\eta = Ce^{\Omega\sigma}\varphi(\sigma) + De^{-\Omega\sigma}\varphi(-\sigma) \tag{1.2}$$

où C et D sont des constantes arbitraires, $\varphi(\sigma)$ une fonction périodique de même periode T que $f(\sigma)$ et Ω une constante déterminée appelée par Poincaré exposant caractéristique. C'est même la solution générale si $\Omega T \neq \pi i$.

La résolution peut se faire par deux méthodes très différentes: la première, par intégration numérique de l'équation; la seconde, due à Hill lui-même, par le calcul d'un déterminant infini.

2. Première méthode. Un théorème de Korteweg [16] est à l'origine du procédé, le résultat a été donné sous une autre forme par Moulton [11] puis étendue par lui [12] au cas d'un système de deux équations du second ordre, donnant les orbites infiniment voisines d'une orbite périodique pour un problème de dynamique; l'application visée avait comme particularité que les fonctions périodiques dans les équations différentielles étaient paires, c'est aussi le cas de l'équation de Hill. On pourrait donc déduire du cas plus général de Moulton, la formule pour l'exposant caractéristique d'une équation de Hill, donnée récemment par Brillouin [2].

Reçu le 27 janvier, 1950.

Nous reprendrons plutôt le raisonnement de façon indépendante, car nous voulons également déterminer la fonction $\varphi(\sigma)$. Supposons déterminées par calcul numérique, deux solutions indépendantes η_1 et η_2, telles que

$$\begin{aligned} &\eta_1(0) = 1,\ \dot\eta_1(0) = 0\text{: solution paire en } \sigma,\\ &\eta_2(0) = 0,\ \dot\eta_2(0) = 1\text{: solution impaire en } \sigma. \end{aligned} \tag{2.1}$$

Le Wronskien des deux solutions, qui est d'ailleurs invariant, a pour valeur l'unité.

La fonction périodique φ peut se décomposer en une fonction paire P et une fonction impaire I, donc, α et β étant des constantes à déterminer

$$\alpha\eta_1 + \beta\eta_2 = e^{\Omega\sigma}(P + I)$$

et en remplaçant σ par $-\sigma$,

$$\alpha\eta_1 - \beta\eta_2 = e^{-\Omega\sigma}(P - I).$$

Si on fait passer l'exponentielle dans le premier membre, on aura en combinant les deux relations:

$$\begin{aligned} P &= \alpha\eta_1 \cosh \Omega\sigma - \beta\eta_2 \sinh \Omega\sigma\\ I &= -\alpha\eta_1 \sinh \Omega\sigma + \beta\eta_2 \cosh \Omega\sigma. \end{aligned} \tag{2.2}$$

Les constantes α et β seront déterminées par la condition de périodicité de P et I; cependant pour que nos écritures soient réelles, nous changerons la définition de Ω dans les cas précisés par la formule (2.4) ci-dessous et écrirons, j valant plus un ou moins un:

$$\begin{aligned} \alpha\eta_1(T) \cosh \Omega T - \beta\eta_2(T) \sinh \Omega T &= j\alpha\eta_1(0) = \alpha j\\ \beta\eta_2(T) \cosh \Omega T - \beta\eta_1(T) \sinh \Omega T &= j\beta\eta_2(0) = 0. \end{aligned} \tag{2.3}$$

Ces équations permettent de déterminer l'exposant caractéristique et le rapport β/α par

$$j \cosh \Omega T = \eta_1(T) \tag{2.4}$$

$$\frac{\beta}{\alpha} = \frac{\eta_1(T)}{\eta_2(T)} \tanh \Omega T. \tag{2.5}$$

Comme la fonction $f(\sigma)$ reprend la même valeur pour des arguments σ et $T - \sigma$, nous pourrons utiliser cette symétrie pour limiter le calcul de η_1 et η_2 à une demi-période et nous pourrons déterminer les $\eta_1(T)$ et $\eta_2(T)$ au moyen des valeurs de ces fonctions pour l'argument $T/2$, que nous symbolisons par (η_1) et (η_2) et des valeurs des dérivées de ces quantités au même endroit: $(\dot\eta_1)$ et $(\dot\eta_2)$.

On trouve en remplaçant dans (2.4) et (2.5):

$$j \cosh \Omega T = (\eta_1)(\dot\eta_2) + (\eta_2)(\dot\eta_1) \tag{2.6}$$

$$\frac{\beta}{\alpha} = \frac{j \sinh \Omega T}{2(\eta_2)(\dot\eta_2)} = \frac{+j\left\{(\eta_1)(\dot\eta_1)(\eta_2)(\dot\eta_2)\right\}^{\frac{1}{2}}}{(\eta_2)\ (\dot\eta_2)}; \tag{2.7}$$

cette dernière égalité parce que

$$\sinh \Omega T = +(\cosh^2 \Omega T - 1)^{\frac{1}{2}} \tag{2.7'}$$

et que le "1" sous le radical peut s'écrire

$$[(\eta_1)(\dot{\eta}_2) - (\eta_2)(\dot{\eta}_1)]^2 \tag{2.8}$$

ce qui représente en effet, le carré du Wronskien invariant à l'endroit de la demi-période.

Seules les valeurs relatives de α et β ont de l'importance, nous décomposerons donc la formule (2.7) en

$$\alpha = (\eta_2)(\dot{\eta}_2)\,|(\eta_2)(\dot{\eta}_2)|^{-\frac{1}{2}} \qquad \beta = j\,|(\eta_1)(\dot{\eta}_1)|^{\frac{1}{2}}. \tag{2.9}$$

Quand le Wronskien ne vaut pas l'unité, il faut employer, au lieu de (2.6)

$$\cosh \Omega T = \frac{(\eta_1)(\dot{\eta}_2) + (\eta_2)(\dot{\eta}_1)}{(\eta_1)(\dot{\eta}_2) - (\eta_2)(\dot{\eta}_1)}\,. \tag{2.10}$$

3. Discussion des résultats. La formule (2.4) nous montre que $j \cosh \Omega T$ peut prendre toutes les valeurs positives et négatives, nous choisirons $j = -1$ si la valeur de $j \cosh \Omega T$ est inférieure à -1.

Si le $\cosh \Omega T$ est plus grand que un en module, on peut trouver des solutions réelles pour Ω et il existe par (1.2) des solutions correspondantes à des orbites voisines de l'orbite périodique considérée, qui s'en éloignent de plus en plus, on dit alors que l'orbite est instable; si $j = +1$ l'instabilité est dite paire [12], expression justifiée par le fait que l'orbite voisine rencontre un nombre pair de fois l'orbite périodique au cours d'une période à cause de (2.2); si $j = -1$, l'instabilité est dite impaire, pour une raison analogue. Dans ce cas les fonctions P et I se développeront en cosinus et sinus impairs de $\omega\sigma/2$; dans chaque cas la solution (1.2) peut se mettre sous la forme

$$A(P \cosh \Omega\sigma + I \sinh \Omega\sigma) + B(P \sinh \Omega\sigma + I \cosh \Omega\sigma). \tag{3.1}$$

Si le $\cosh \Omega T$ est en module inférieur à un, il existe des solutions de (2.4) qui sont des imaginaires purs, nous écrirons donc $\Omega = \Omega' i$; les formules ci-dessus seront rendues réelles, si on pose même temps, $P = P'$, $I = I' i$, α et β ayant la même définition qu'en (2.9); on aurait alors

$$\begin{aligned} P' &= \alpha\eta_1 \cos \Omega'\sigma + \beta\eta_2 \sin \Omega'\sigma \\ I' &= -\alpha\eta_1 \sin \Omega'\sigma + \beta\eta_2 \cos \Omega'\sigma. \end{aligned}$$

On peut prendre cette fois à son gré l'une ou l'autre détermination pour Ω' par

$$j \cos \Omega' T = (\eta_1)(\dot{\eta}_2) + (\eta_2)(\dot{\eta}_1),$$

les développements se faisant en sinus et cosinus soit pairs soit impairs de $\omega\sigma/2$ et la solution générale peut cette fois se mettre sous une forme équivalente à (1.2):

(3.2) $$A'(P' \cos \Omega'\sigma - I' \sin \Omega'\sigma) + B'(P' \sin \Omega'\sigma + I' \cos \Omega'\sigma).$$

Nous donnons une application de ces formules au N° 11.

4. Examen des cas limites. Il nous reste à examiner maintenant ce qui se passe lorsque le $\cosh \Omega T = \pm 1$.

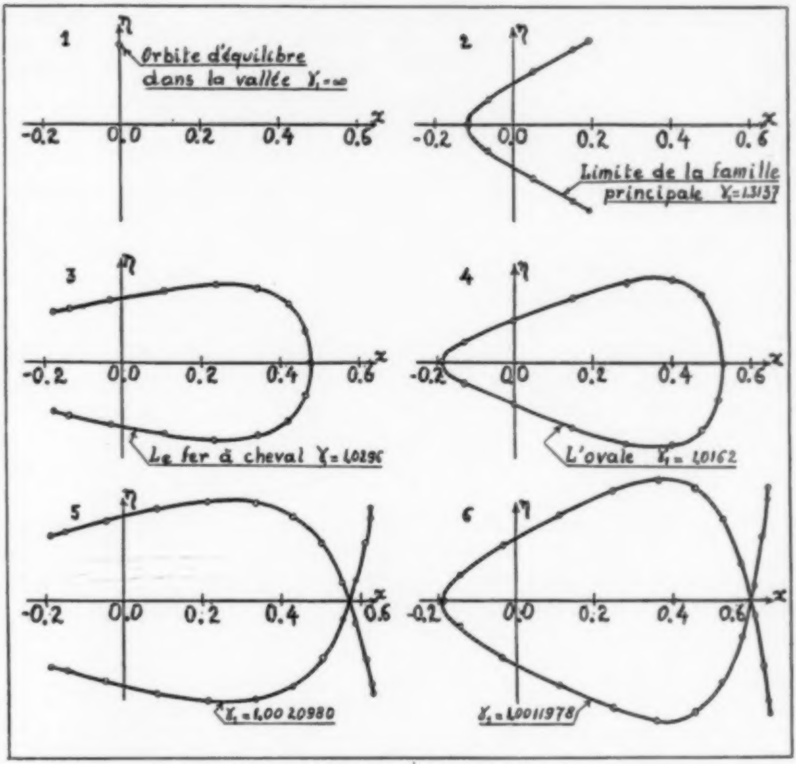

FIGURE 1. Orbites périodiques infiniment voisines de l'équateur. (L'échelle des η est arbitraire.)

(i): $\cosh \Omega T = +1$. En comparant (2.6) et (2.8) on voit que $(\eta_2)(\dot{\eta}_1) = 0$.

Si $(\eta_2) = 0$, $\alpha = 0$ seule la solution impaire I subsiste; on a une orbite qui rencontre $\lambda = 0$ en ses deux extrémités comme l'orbite 4 de la figure 1, en supposant $\sigma = 0$ pour x minimum.

Si $(\dot{\eta}_1) = 0$, $\beta = 0$, seule la solution paire subsiste; on a une orbite dont la vitesse s'annule aux deux extrémités comme l'orbite 5 de la figure 1.

(ii): $\cosh \Omega T = -1$. Cette fois on a $(\eta_1)(\dot{\eta}_2) = 0$.

Si $(\eta_1) = 0$, $\beta = 0$, on a une orbite comme celle 3 de la figure 1.
Si $(\dot{\eta}_2) = 0$, $\alpha = 0$, on a une orbite comme celle 2 de la figure 1.

5. Deuxième méthode, généralités. La méthode de Hill consiste à développer en séries, la fonction périodique $f(\sigma)$ et la fonction inconnue $\varphi(\sigma)$ intervenant dans la solution; en remplaçant dans l'équation (1.1), on obtient une infinité de relations entre les coefficients de la série $\varphi(\sigma)$; la compatibilité de ces relations s'exprime au moyen d'un déterminant infini égalé à zéro, ce qui permet de déterminer l'exposant caractéristique Ω, soit par approximations successives [3] et [8], soit par un développement en série [3, pp. 35-37].

Nous n'utiliserons la méthode de Hill que pour améliorer une première approximation de la solution et cela dans le seul cas des orbites limites.

6. Amélioration des orbites limites. Nous avons vu que pour les orbites limites, l'exposant caractéristique est nul et la fonction $\varphi(\sigma)$ se réduit à sa partie paire ou impaire; nous exposerons la méthode d'amélioration dans le cas où $\varphi(\sigma)$ se développe en sinus de multiples impairs de $\omega\sigma/2$; les autres se traiteront par analogie.

Nous traitons donc le cas où $f(x)$ dépend aussi d'un paramètre, disons γ; nous supposerons connaître une valeur γ_0 de γ, où l'exposant caractéristique est presque nul et cherchons à améliorer cette valeur; nous nous y prenons comme suit: nous faisons correspondre à l'équation de Hill (1.1), l'équation

$$A \frac{d^2\eta}{d\sigma^2} + f(\sigma)\eta = 0 \; ; \tag{6.1}$$

celle-ci admet pour certaines valeurs de A une solution périodique, pour une orbite limite vérifiant (1.1), $A = 1$ est une telle valeur; nous devons donc nous attendre que pour une valeur de γ donnant un exposant caractéristique de (1.1) presque nul, on aura une solution périodique pour A voisin de un.

Nous écrivons d'abord, $2\pi/\omega$ étant la période de $f(\sigma)$:

$$4f(\sigma)/\omega^2 = B_0 + \sum_{m=1}^{\infty} 2\, B_m \cos m\omega\sigma, \tag{6.2}$$

puis la solution périodique de (6.1)

$$\eta = \sum_{p=1}^{\infty} S_p \sin (2p - 1)\, \omega\sigma/2. \tag{6.3}$$

En remplaçant dans (6.1) nous avons le système suivant à résoudre:

$$\begin{aligned} (B_0 - B_1)S_1 + (B_1 - B_2)S_2 + (B_2 - B_3)S_3 + \dots &= AS_1, \\ (B_1 - B_2)S_1 + (B_0 - B_3)S_2 + (B_1 - B_4)S_3 + \dots &= AS_2, \\ (B_2 - B_3)S_1 + (B_1 - B_4)S_2 + (B_0 - B_5)S_3 + \dots &= AS_3, \end{aligned} \tag{6.4}$$

dont nous sommes supposés connaître une approximation pour S_p et pour A (un). η n'est déterminé qu'à une constante multiplicative près, nous devons donc fixer la valeur d'une des harmoniques par exemple S_1.

Un tel système peut se résoudre par itération en portant les approximations des S_p dans le premier membre; la première équation donne A, les suivantes une nouvelle approximation des S_p.

Cette méthode n'est cependant efficace que si la première harmonique prédomine; (les exemples du N° 11 sont dans ce cas) dans les cas contraires on pourra procéder comme suit, si les q premières harmoniques prédominent: décomposons le système (6.4) en deux systèmes partiels, (a) constitué par les q premières équations et (b) pas les dernières; on commence par résoudre (b) par itération en supposant $A = 1$ et en se donnant pour les q premières harmoniques leur première approximation; les autres harmoniques ont leurs valeurs qui convergent rapidement vers une deuxième approximation qu'on porte dans le système (a), celui-ci se résout par la formule de Cramer en prenant comme inconnues les corrections des S et A, d'où une deuxième approximation des q premières harmoniques; on recommence alors à traiter (b) puis (a). Le calcul des mineurs normés de (a) ne doit se faire qu'une fois, car la correction en passant d'une approximation à la suivante est petite, ceci est un sérieux avantage du procédé.

Nous allons maintenant donner des applications de ce deuxième procédé au problème de Störmer (N^{os} 11 et 17).

Orbites Equatoriales du Problème de Störmer

7. Généralités. L'etude du mouvement d'une particule électrisée dans le champ d'un dipôle ou problème de Störmer, se réduit [7] et [15] à celle du mouvement dans le plan méridien qui suit la particule et à celle du mouvement du plan méridien.

Prenant comme coordonnées $x = \log 2\gamma_1 r$ et λ, où r est la distance au dipôle, λ la latitude et γ_1 un paramètre lié au mouvement du plan méridien et qu'on remplace aussi par $a = 1/(16\,\gamma_1^4)$, on trouve

$$\ddot{x} = ae^{2x} - e^{-x} + e^{-2x}\cos^2\lambda \tag{7.1}$$

$$\ddot{\lambda} = -(1 + \mathrm{tg}^2\lambda - e^{-2x}\cos^2\lambda)\,\mathrm{tg}\lambda \tag{7.2}$$

équations admettant comme intégrale première

$$\dot{x}^2 + \dot{\lambda}^2 = ae^{2x} - 1 - \mathrm{tg}^2\lambda + 2e^{-x} - e^{-2x}\cos^2\lambda. \tag{7.3}$$

Parmi les trajectoires satisfaisant à ces équations, celles qui sont périodiques méritent une étude spéciale, soit pour augmenter notre connaissance de la théorie des orbites dans les problèmes non intégrables de la dynamique, soit pour préparer le calcul des cônes du rayonnement cosmiques.

Diverses familles d'orbites périodiques ont été découvertes et calculées [10], [13] et [15]; nous nous proposons de résumer ici ce qui a été fait pour les orbites sur l'équateur $\lambda = 0$ et de compléter cette étude par le calcul des exposants caractéristiques de ces orbites.

8. Résultats connus. Résumons d'abord pour clarifier les idées comment se présentent ces orbites périodiques. Tout d'abord, elles n'existent que pour des valeurs de γ_1 variant de un à l'infini; pour des valeurs infiniment grandes, les orbites se réduisent au point $x = 0$; lorsque γ_1 diminue, les orbites oscillent entre des points qui s'éloignent de part et d'autre de $x = 0$, jusqu'à atteindre pour $\gamma_1 = 1$ les limites $x = \log 2\ (2^{\frac{1}{2}} - 1)$ et $x = \log 2$, mais cette dernière orbite ne peut plus être strictement dite périodique car sa période devient infinie.

Une première étude de ces orbites pour des valeurs de γ_1 voisines de un a été faite par G. Lemaître [6]; il y est prouvé une relation approchée existant entre le paramètre $\alpha = 16a - 1$ et la demi-période σ_m:

$$\sigma_m = -2^{-\frac{1}{2}} \log(-\alpha/64). \tag{8.1}$$

Nous y reviendrons dans la suite.

Dans une autre étude faite par C. Graef et S. Kusaka, [4] on a déterminé le mouvement dans l'espace correspondant à ces orbites.

Mais l'étude la plus intéressante est certes celle de la manière dont se présente le voisinage de ces orbites périodiques si on ne reste plus sur l'équateur $\lambda = 0$. G. Lemaître a résolu le problème pour des valeurs de γ_1 voisines de un dans un travail non publié qu'il nous a permis de reprendre ici. Il a été amorcé aussi par J. Lifshitz [9].

Nous avons nous-mêmes, en annexe à notre thèse de doctorat, fait certaines déterminations et nous venons d'épuiser le problème.

9. Equations aux variations. Les orbites infiniment voisines de l'équateur de coordonnées $x + \xi$ et η se détermineront au moyen des solutions des équations aux variations déduites de (7.1 à 7.3), ce sont

$$\ddot{\xi} = (2ae^{2x} - 2e^{-2x} + e^{-x})\,\xi, \tag{9.1}$$

$$\ddot{\eta} = (e^{-2x} - 1)\,\eta, \tag{9.2}$$

$$\dot{x}\dot{\xi} = \ddot{x}\xi. \tag{9.3}$$

On remarque que les variables ξ et η sont indépendantes; de plus, l'équation (9.3) s'intègre immédiatement et donne $\xi = C\dot{x}$, où C est arbitraire; mais cette solution est banale car $x + C\dot{x}$, représente simplement l'orbite sur l'équateur parcourue un peu plus tard. On voit donc, qu'en première approximation, il ne faut pas considérer les variations de x, contrairement à ce qui est fait dans l'article [9].

Le problème revient donc à résoudre les équations (7.1) et (9.2); c'est ce que nous ferons d'abord pour γ_1 suffisamment différent de un, puis nous verrons comment utiliser un travail de Tchang Yong-Li [14] pour les valeurs de γ_1 voisines de un, enfin nous verrons comment améliorer cette approximation quand cela est nécessaire.

VALEURS DE γ_1 NON VOISINES DE UN

10. Calcul de l'orbite sur l'équateur. Plusieurs méthodes peuvent être utilisées pour résoudre l'équation (7.1) ou l'équation équivalente (7.3) quand $\lambda = 0$. Dans chacune on utilise les propriétés des fonctions elliptiques. Nous emploierons d'abord la suivante utilisée depuis longtemps au département de mathématiques de l'Université de Louvain.

En posant

$$a_1^2 = \tfrac{1}{4}(1 + \gamma_1^{-2}) \tag{10.1}$$

$$b_1^2 = \tfrac{1}{4}(1 - \gamma_1^{-2}) \tag{10.2}$$

l'équation (7.3) peut s'écrire

$$\left(\frac{de^{-x}}{d\sigma}\right)^2 = [a_1^2 - (e^{-x} - \tfrac{1}{2})^2]\,[(e^{-x} - \tfrac{1}{2})^2 - b_1^2] \tag{10.3}$$

et la fonction dn d'Abel et de Jacobi permet d'écrire immédiatement la solution

$$e^{-x} = \tfrac{1}{2} + a_1 \operatorname{dn} a_1\sigma = \tfrac{1}{2} + a_1 \operatorname{dn} u. \tag{10.4}$$

Ceci peut se calculer au moyen des fonctions de Jacobi:

$$\operatorname{dn} u = \frac{\theta_3(u)\ \theta_2(0)}{\theta_2(u)\ \theta_3(0)}$$

où on a avec $2\,\omega_1 s = \pi u$

$$\theta_2(u) = 1 - 2q\cos 2s + 2q^4\cos 4s - 2q^9\cos 6s + \ldots$$
$$\theta_3(u) = 1 + 2q\cos 2s + 2q^4\cos 4s + 2q^9\cos 6s + \ldots.$$

En vertu des relations (10.1) et (10.2) et de

$$\frac{b_1}{a_1} = \left[\frac{\theta_2(0)}{\theta_3(0)}\right]^2$$

l'arbitraire "q" est liée à la constante γ_1 par la relation

$$\gamma_1^2 = \frac{\theta_3^4(0) + \theta_2^4(0)}{\theta_3^4(0) - \theta_2^4(0)},$$

on trouve aussi pour la vitesse moyenne

$$\omega = 2\,\pi/T = 2^{\frac{1}{2}}\{\theta_3^4(0) + \theta_2^4(0)\}^{-\frac{1}{2}},$$

tandis que

$$2a_1^2 = \frac{\theta_3^4(0)}{\theta_3^4(0) + \theta_2^4(0)}.$$

11. Calculs et résultats. Nous avons calculé par les formules précédentes un certain nombre d'orbites sur l'équateur; les résultats sont condensés au

TABLEAU I

ORBITES ÉQUATORIALES, LEURS EXPOSANTS CARACTÉRISTIQUES, LEURS VARIATIONS

γ_1	∞	2.5100	1.7957	1.4939	1.3166	1.3137	1.1394	1.0882	1.0304	1.0236
ω	1.0000	0.9952	0.9813	0.9593	0.9284	0.9275	0.8509	0.8024	0.6933	0.6698
q	0.0000	0.02000	0.04000	0.06000	0.08147	0.08197	0.12500	0.15000	0.20700	0.22000
$\cosh \Omega T$	1.0000	0.8584	0.4529	−0.1679	−0.9800	−1.0000	−2.4833	−2.8193	−1.0794	−0.2655
α	1.7725	1.8840	1.9483	1.9590	1.9116	1.9100	1.6152	1.3122	0.1373	−0.3178
β	0.0000	0.1360	0.2288	0.1402	0.0052	0.0000	−0.7037	−1.0045	−1.4764	−1.5116
x_0	0.0000	0.00280	0.01097	0.02400	0.04262	0.04307	0.09039	0.12127	0.1933	0.20929
x_1		−0.03989	−0.07912	−0.11707	−0.15581	−0.15667	−0.22596	−0.26044	−0.3222	−0.33310
x_2		−0.00040	−0.00161	−0.00362	−0.00671	−0.00679	−0.01592	−0.02298	−0.0436	−0.04904
x_3			−0.00004	−0.00014	−0.00037	−0.00037	−0.00131	−0.00227	−0.0059	−0.00722
x_4					−0.00002	−0.00003	−0.00012	−0.00025	−0.0009	−0.00118
x_5							−0.00001	−0.00003	−0.0002	−0.00020
x_6										−0.00004
x_7										−0.00001
		st.	st.	st.	st.		imp.	imp.	imp.	st.
P_0	3.5449	3.4809	3.2978	2.9381	1.9641	1.7894				1.3077
P_2		0.2863	0.5938	0.9611	1.7603	1.9118				−1.3734
P_4		0.0010	0.0050	0.0192	0.1004	0.1206				−0.5809
P_6				−0.0003	−0.0017	−0.0018				0.0112
P_8										−0.0002
I_2	0.0000	−0.0483	−0.2065	−0.5196	−1.4737	−1.6669				3.4186
I_4		−0.0007	−0.0051	−0.0209	−0.1045	−0.1245				0.6775
I_6				0.0004	0.0018	0.0019				−0.0219
I_8										0.0010
P_1	3.5449	3.6479	3.6981	3.6786	3.5811	3.5787	3.1105	2.6403	0.3788	−1.0883
P_3		0.1200	0.1988	0.2408	0.2458	0.2450	0.1328	0.0048	−0.0909	0.3934
P_5				−0.0014	−0.0038	−0.0038	−0.0138	−0.0224	−0.0147	0.0648
P_7					0.0001	0.0001	0.0009	0.0017	0.0014	−0.0063
P_9										0.0008
I_1	−3.5449	−3.3136	−2.8976	−2.1978	−0.3470	0.0000	−3.2614	−3.9857	−4.0984	−3.7036
I_3		−0.1182	−0.1889	−0.2008	−0.0409		−0.5637	−0.8383	−1.4881	−1.6518
I_5				0.0004	0.0003		0.0087	0.0155	0.0009	−0.0318
I_7								−0.0003	0.0020	−0.0044
I_9										−0.0007

Tableau I où toutes les quantités sont indiquées en unités de la quatrième décimale. On y voit les analyses harmoniques des orbites données par

$$x = \sum_{k=0}^{n} x_k \cos k\omega\sigma.$$

La solution des équations aux variations a été faite par le procédé des N^{os} 2 et 3 *en prenant comme point de départ le point de l'orbite ayant un x maximum*, contrairement à ce qui est fait dans le reste du travail. Nous avons

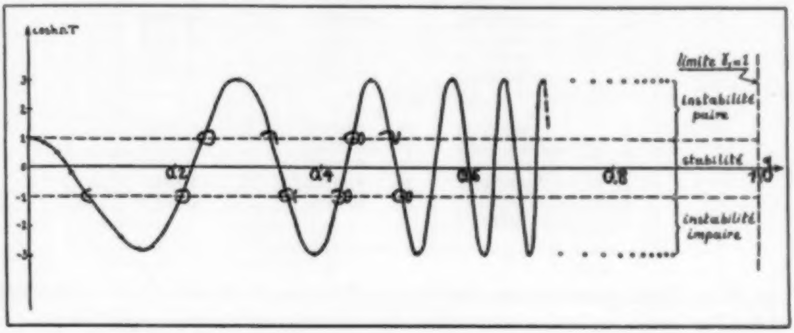

FIGURE 2. Exposants caracteristiques des orbites périodiques sur l'équateur.

indiqué si les orbites sont stables (st.) ou instables impaires (imp.), ainsi que les analyses en série de Fourier des fonctions P et I des solutions (3.1) et (3.2) par

$$P = \sum_{k=0}^{n} P_k \cos \frac{k\omega\sigma}{2} \text{ et } I = \sum_{k=0}^{n} I_k \sin \frac{k\omega\sigma}{2}.$$

Dans le cas d'orbites stables, nous avons donné d'abord le résultat avec $j = 1$ (harmoniques paires), puis avec $j = -1$ (harmoniques impaires), une seule forme suffit évidemment; dans le cas d'orbites instables impaires, on doit poser (N° 3) $j = -1$ et les résultats se développent avec des harmoniques impaires de $\omega\sigma/2$.

Nous avons aussi donné les valeurs limites pour q tendant vers zéro ($2\pi^{\frac{1}{2}} = 3.5449$) qui peuvent servir pour trouver par interpolation toute orbite dans l'intervalle du Tableau I.

On peut constater par les valeurs du cosh ΩT que si q augmente à partir de zéro, les orbites de l'équateur sont successivement stables, instables impaires et stables; la première partie de la figure 2 a été construite au moyen de ces résultats.

Il existe donc des orbites limites que nous avons déterminées par approximations successives avec la méthode du N° 6. La forme des orbites limites pouvait cependant être prévue pas l'examen du Tableau I, grâce à la discussion

TABLEAU II

QUELQUES ORBITES LIMITES

($\sigma = 0$ pour le point de l'orbite dont l'abscisse x est minimum)

q	0.08198063	0.20835	0.2388
γ_1	1.3135943	1.0296	1.0162
ω^2	0.8602794	0.4772	0.4057
x_0	0.04307791	0.1949	0.2319
x_1	−0.15669039	−0.3234	−0.3468
x_2	−0.00679391	−0.0441	−0.0574
x_3	−0.00036765	−0.0061	−0.0092
x_4	−0.00002259	−0.0010	−0.0016
x_5	−0.00000149	−0.0002	−0.0003
x_6	−0.00000010	−0.0001	−0.0001
	$\eta = I$	$\eta = P$	$\eta = I$
I_1	1.00000000	P_1 1.0000	I_2 1.0000
I_3	−0.06860992	P_3 −0.3700	I_4 −0.1645
I_5	−0.00109447	P_5 0.0002	I_6 −0.0104
I_7	−0.00003392	P_7 0.0004	I_8 −0.0011
I_9	−0.00000121	P_9 0.0001	I_{10} −0.0002
I_{11}	−0.00000004		

du N° 4. Nous donnons ces résultats au Tableau II et à la figure 1 (orbites 1 à 4.)

Nous n'avons pas continué au-delà de $q = 0.2200$; d'abord les calculs deviennent plus long, car les séries de Fourier ont une convergence pratique moins bonne, ensuite une approximation suffisante peut être obtenue plus aisément par le procédé que nous indiquons maintenant.

VALEURS DE γ_1 VOISINES DE UN

12. Généralités. La solution des équations au voisinage de l'équateur et pour des valeurs de γ_1 suffisamment rapprochées de un a été calculée par L. Bouckaert [1] et T. Yong-Li [14]. Une première approximation s'écrit avec $a = 1/\gamma_1^4 - 1$:

$$x = \log 2 - \log (1 + 2^{\frac{1}{2}} \operatorname{sech} \Omega\sigma) + U_a \tag{12.1}$$

$$\eta = A_1 \sin (\omega\sigma + \varphi_0) + A_2 \cos (\omega\sigma + \varphi_0) = A \sin (\omega\sigma + \varphi_0 + a) \tag{12.2}$$

où $\omega = \frac{1}{2} 3^{\frac{1}{2}}$, $\Omega = \frac{1}{2} 2^{\frac{1}{2}}$ et U_a, A_1, A_2, A et a sont des fonctions tabulées pour les valeurs négatives de σ, telles que $\sigma = 0$ pour le point de l'orbite le plus proche du dipôle (rejeté dans cette représentation à l'infini négatif) et telles que $A(-\infty) = 1$ et $a(-\infty) = 0$.

Nous reprendrons dans les deux numéros suivants le travail de G. Lemaître, qui permet de déterminer une approximation de l'exposant caractéristique des orbites et aussi une approximation des orbites limites; nous commencerons par ces dernières. Nous ne chercherons pas à ajouter de la précision en considérant les termes suivants du développement de Tchang Yong-Li, car l'approximation

TABLEAU III

COMPARAISON DES RÉSULTATS APPROCHÉS ET EXACTS

POUR QUELQUES ORBITES ÉQUATORIALES REMARQUABLES

$\delta = -\alpha/4$	$\epsilon = \gamma_1 - 1$		q	$\cosh \Omega T$
	approché	exact	exact	
		∞	0.00000	+1
0.02014	0.2531	0.3136	0.08198	−1
0.0738	0.0913			min
0.0270	0.0290	0.0296	0.29835	−1
0.0155	0.0161	0.0162	0.2388	+1
0.005674	0.005755			max
0.002078	0.002089	0.0020980	0.33152	+1
0.0011914	0.0011950	0.0011978	0.35382	−1
0.00043636	0.00043684			min
0.00015982	0.00015988		0.4244*	−1
0.00009163	0.00009165		0.4415	+1
0.000033561	0.000033564		0.4701	max
0.000012292	0.000012292		0.49612	+1
0.000007047	0.000007047		0.50949	−1
0.0000025812	0.0000025812		0.53203	min
0.0000009454	0.0000009454		0.55268	−1
0.0000005420	0.0000005420		0.56348	+1

* *à partir de cet endroit q a été calculé par la formule (16.6).*

est suffisante et les calculs deviennent sinon beaucoup plus compliqué, tandis que l'amélioration pourra se faire aisément comme nous l'indiquerons plus loin (N° 16).

13. Orbites limites. Pour qu'une orbite infiniment voisine de l'équateur soit périodique il suffit:

1°) Qu'au départ ($\sigma = 0$) la fonction λ soit paire ou impaire;

paire si $\dot{\lambda} = 0$, ce qui implique $\varphi_0 = \varphi_i = -a(0)$;
impaire si $\lambda = 0$, ce qui implique $\varphi_0 = \varphi_p = -a(0) - 2\chi$;

les quantités $a(0) = -63°\,30'\,10''$ et $\chi = 9°\,45'\,31''$ ont été déterminées avec cette précision par T. Yong-Li [14].

2°) Il faut qu'il en soit de même à l'autre extremum de x, $\sigma = -\sigma_m$; à cet endroit nous supposerons que σ_m est suffisamment grand pour que $A = 1$ et $a = 0$, donc

$$\eta = \sin(-\omega\sigma_m + \varphi_0) \text{ et } \dot{\eta} = \omega \cos(-\omega\sigma_m + \varphi_0);$$

il suffira donc que

$$\omega\sigma_m - \varphi_0 = k\pi/2,$$

si k est pair $\eta = 0$ et si k est impair $\dot{\eta} = 0$.

TABLEAU IV

Résumé des différents stades d'approximation pour
$\epsilon = 0.0011960$

B_m	approxim. de départ		(b)	(a)	(b)	(a)	(b)
7.53329	A	1.00000		1.00048		1.00044	
−4.31112	S_1	−0.25000		−0.25000		−0.25000	
−2.07372	S_2	−1.40564		−1.38768		−1.38762	
−0.91168	S_3	0.32424		0.32214		0.32232	
−0.38352	S_4	0.03617	0.03423		0.03361		0.03359
−0.15856	S_5	0.00440	0.00587		0.00574		0.00575
−0.06396	S_6	0.00239	0.00123		0.00120		0.00120
−0.02540	S_7	−0.00077	0.00028		0.00027		0.00027
−0.00996	S_8	0.00097	0.00007		0.00007		0.00007
−0.00388	S_9	0.00000	0.00002		0.00002		0.00002
−0.00152							
−0.00056							
−0.00024							
−0.00008							

Ceci peut être combiné avec l'expression de φ_0 d'après le 1° et avec celle de σ_m tirée de (8.1). Ceci donne les deux formules

$$\text{(13.1)} \quad \delta = -a/4 = 16e^{4a(0)/\epsilon^{\frac{1}{2}}}\, e^{-2k\pi/\epsilon^{\frac{1}{2}}} = 2.6187(0.076911)^k$$

$$\text{(13.2)} \quad \delta = -a/4 = 16e^{4a(0)/\epsilon^{\frac{1}{2}}}\, e^{-2k\pi/\epsilon^{\frac{1}{2}}}\, e^{8\chi/\epsilon^{\frac{1}{2}}} = 4.5674(0.076911)^k$$

en comparant ce qui vient d'être écrit avec le N° 4 on verra que pour (13.1) et k pair on a une orbite du type c du N° 4, si k est impair c'est une orbite du type b; pour (13.2) on aura pour k pair le type a et k impair le type d.

Nous avons résumé dans le Tableau III les valeurs δ ainsi obtenues, d'ailleurs comme

$$\text{(13.3)} \qquad \gamma_1 = (1 - 4\delta)^{-\frac{1}{4}} = 1 + \delta + \tfrac{5}{2}\,\delta^2 + \tfrac{15}{2}\,\delta^3 + \dots$$

on peut dire qu'au premier ordre, et c'est à cet ordre que nous avons limité les développements de Tchang Yong-Li, $\gamma_1 - 1 = \delta$; cependant si nous calculons γ_1 avec la formule exacte (13.3), il se fait que ces valeurs sont toujours plus proches des valeurs réelles calculées d'après le N° 11 et que nous comparons dans le même tableau.

Nous proposons donc cette variante. Il nous reste maintenant à calculer les exposants caractéristiques aux points intermédiares.

14. Exposants caractéristiques. Nous avons vu au 1° et 2° du numéro précédent comment se calculent les solutions qui pour $\sigma = 0$ sont paires ou impaires et comment se calculent les fonctions et leurs dérivées pour $\sigma = -\sigma_m$; en remplaçant dans (2.10) et en réduisant on trouve la formule simple

$$\text{(14.1)} \qquad \cosh \Omega T = \frac{-\sin\,[2\,\omega\sigma_m + 2a(0) + 2\chi]}{\sin 2\chi}$$

TABLEAU V

DEUX AUTRES ORBITES LIMITES

γ_1 q ω	1.0011978 0.35382 0.46718	1.0020980 0.33152 0.49622	
$B_m/4$	$\eta(S_p = I_{2p-1})$	$B_m/4$	η
1.88247	I_1 −0.25000	1.58328	P_0 0.44212
−1.07748	I_3 −1.38803	−0.97439	P_2 1.00000
−0.51820	I_5 0.32247	−0.44134	P_4 −0.61869
−0.22779	I_7 0.03358	−0.18278	P_6 −0.00758
−0.09626	I_9 0.00575	−0.07266	P_8 0.00079
−0.03960	I_{11} 0.00120	−0.02808	P_{10} 0.00042
−0.01597	I_{13} 0.00028	−0.01064	P_{12} 0.00014
−0.00634	I_{15} 0.00007	−0.00397	P_{14} 0.00005
−0.00249	I_{17} 0.00002	−0.00146	P_{16} 0.00001
−0.00097		−0.00053	
−0.00038		−0.00019	
−0.00014		−0.00007	
−0.00006		−0.00003	
−0.00002		−0.00001	

en égalant à ± 1, on retrouve les orbites limites ci-dessus, avec pour a et b, $\cosh \Omega T = 1$ et pour c et d, $\cosh \Omega T = -1$.

On voit d'autre part que pour γ_1 suffisamment voisin de un le $\cosh \Omega T$ varie sinusoïdalement (voir figure 2) avec des maxima et minima valant en module $1/\sin 2\chi = 2.9927$. Ces extrema ont lieu pour

$$\delta = 16e^{4a(0)/q^{\frac{1}{2}}}\, e^{(2k-1)\pi/q^{\frac{1}{2}}}\, e^{4\chi/q^{\frac{1}{2}}} = 9.5913(0.076911)^k$$

ils sont aussi indiqués au Tableau III.

15. Comparaisons des résultats. La comparaison des résultats exacts du N° 11 et approchés du N° 13 montre que trois décimales sont déjà bonnes pour $\gamma_1 = 1.0296$ et que cela s'améliore encore; on voit donc l'utilité de l'approximation de T. Yong-Li et comment se succèdent les zones de stabilité et d'instabilité; comme pratiquement seule une courbe approchée de l'exposant caractéristique est nécessaire dès qu'on connaît les orbites limites, il nous suffira de montrer qu'on peut améliorer comme on le désire toutes les orbites limites. On aura ainsi épuisé le problème posé par la famille équatoriale. Nous tenons cependant à remarquer ici que d'autres familles d'orbites périodiques existent au voisinage de l'équateur, lorsque l'orbite équatoriale est stable et que $\Omega' = -i\Omega$ est commensurable avec ω; mais ceci est un autre problème.

AMELIORATION DES ORBITES LIMITES

16. Variante du calcul de l'orbite. Lorsque q augmente, le calcul de l'orbite par les fonctions θ devient de moins en moins aisé. Il existe alors une autre méthode qui a été exposée en détail par Lifshitz [9] dont le principe est

de déterminer théoriquement les coefficients du développement en série de Fourier de la fonction périodique $e^{-2x} - 1$ de l'équation de Hill (9.2). Lifshitz utilisait une autre transformation que celle du N° 10 pour exprimer e^{-x}. Nous avons repris ses calculs avec la transformation (10.4).

On trouve dans Whittaker et Watson [17] aux N° 22.6 ex. 1 et 22.735 ex. 5, des formules qui peuvent s'écrire en posant à cause de (10.4).

$$x = \frac{u\pi}{K} = \frac{a_1\pi}{K}\sigma = \omega\sigma \tag{16.1}$$

$$\operatorname{dn} u = \frac{\pi}{2K} + \frac{2\pi}{K}\sum_{m=1}^{\infty}\frac{q^m \cos m\omega\sigma}{1+q^{2m}}$$

$$\operatorname{dn}^2 u = 1 - k^2 \operatorname{sn}^2 u = \frac{E}{K} + \frac{2\pi^2}{K^2}\sum_{m=1}^{\infty}\frac{mq^m \cos m\omega\sigma}{1-q^{2m}}.$$

On en déduit les coefficients a_m de la série de Fourier représentant $f(\sigma)$:

$$f(\sigma) = (\tfrac{1}{2} + a_1 \operatorname{dn} u)^2 - 1 = \sum_{m=0}^{\infty} a_m \cos m\omega\sigma$$

et donc les coefficients B_m qui y sont rattachés par (6.2):

$$B_0 = 3/\omega^2 - 2/\omega - 4EK/\pi^2 \tag{16.2}$$

$$B_m = -\frac{4mq^m}{1-q^{2m}} - \frac{4q^m}{\omega(1+q^{2m})}. \tag{16.3}$$

(Le ω qui intervient dans ces formules vaut le double de celui de Lifshitz.)

Remarquons incidemment que la transformation (10.4) pourrait se déduire de celle de Lifshitz par une transformation de Landen. On pourrait calculer les intégrales complètes K, E et K' par une ou plusieurs transformations successives; mais lorsque γ_1 est plus petit que 1.002, pour une précision de cinq décimales, les termes des développements suivants en $\epsilon = 1 - \gamma_1$ sont négligeables à partir du second order en ϵ :

$$\begin{aligned}
a_1 &= 2^{-\frac{1}{2}}\left(1 - \tfrac{1}{2}\epsilon + \tfrac{5}{8}\epsilon^2 + \dots\right),\\
k'^2 &= \epsilon - \tfrac{1}{2}\epsilon^2 + \dots,\\
K' &= \tfrac{1}{2}\pi\left(1 + \tfrac{1}{4}k'^2 + \tfrac{9}{64}k'^4 + \dots\right),\\
K &= \left(1 + \tfrac{1}{4}k'^2 + \tfrac{9}{64}k'^4 + \dots\right)\log\frac{4}{k'} - \tfrac{1}{4}k'^2 - \tfrac{21}{128}k'^4 - \dots,\\
E &= \left(\tfrac{1}{2}k'^2 + \tfrac{3}{16}k'^4 + \dots\right)\log\frac{4}{k'} + 1 - \tfrac{1}{4}k'^2 - \tfrac{13}{64}k'^4 - \dots,\\
q &= e^{-\pi K'/K}.
\end{aligned} \tag{16.4}$$

On a donc, à cette approximation et en écrivant $\epsilon_1 = \frac{1}{4}\,\epsilon$:

$$
\begin{aligned}
a_1 &= 2^{-\frac{1}{2}}(1 - 2\epsilon_1), \\
k'^2 &= 4\epsilon_1, \\
K' &= \tfrac{1}{2}\,\pi(1 + \epsilon_1), \qquad (16.5) \\
K &= (1 + \epsilon_1) \log\,(2\epsilon_1^{-\frac{1}{2}}) - \epsilon_1, \\
E &= 2\epsilon_1 \log\,(2\epsilon_1^{-\frac{1}{2}}) + 1 - \epsilon_1.
\end{aligned}
$$

Enfin, si on se limite au premier terme de chaque développement on retrouve d'une part la formule (8.1):

$$\sigma_m = \frac{K}{a_1} = -\,2^{-\frac{1}{2}} \log \left(\frac{k'}{4}\right)^2 = -\,2^{-\frac{1}{2}} \log \left(-\,\frac{a}{64}\right)$$

et d'autre part la formule

$$\log q \log\,(\epsilon/16) = \pi^2 \qquad (16.6)$$

qui nous permet de calculer q pour les valeurs de ϵ suffisamment petites, afin de compléter la dernière partie de la figure 2.

17. Calculs et résultats. Nous avons cherché à améliorer les orbites limites dont une première approximation de ϵ était donnée au Tableau III à savoir $\epsilon = 0.002089$ et 0.001195. Nous avons d'abord déterminé l'orbite sur l'équateur qui correspond à ces valeurs de ϵ par la méthode du numéro précédent, en particulier les harmoniques B_m intervenant dans la formule (6.2) sont données par (16.2) et (16.3). Puis nous avons déterminé une première approximation de la solution au moyen de la formule (12.1) ci-dessus, de la formule (7) du travail de Tchang Yong-Li [14] et de son Tableau I transformé en série de puissances de $\cos\theta$.

Les approximations successives sont déterminée par la méthode du N° 6, avec $q = 3$ et dont nous donnons un exemple pour $\epsilon = 0.0011960$ au Tableau IV, la dernière approximation nous donne aussi la valeur de A: 1.00044. Nous avons recommencé pour $\epsilon = 0.0011980$; A vaut alors 0.99995, par interpolation linéaire on trouvera qu'à cinq décimales l'orbite limite cherchée a lieu pour $\gamma_1 = 1.0011978$.

Nous avons résumé au Tableau V la solution pour cette orbite; celle qui correspond à $\gamma_1 = 1.0020980$ a été obtenue de façon semblable.

18. Conclusions. Au sujet des problèmes où interviennent la résolution d'une équation de Hill, nous avons mis au point (N° 2 et 3) une méthode permettant de calculer par intégration numérique non seulement l'exposant caractéristique, mais aussi la solution de l'équation. Nous avons également montré (N° 6) comment trouver par approximations successives, une orbite limite (d'exposant caractéristique nul) dès qu'on connaît une première approximation.

Pour le problème de Störmer, nous avons calculé les exposants caractéristiques de la famille d'orbites périodiques sur l'équateur (fig. 2) ainsi que quelques orbites limites (fig. 1).

Ce travail montre comment une infinité d'orbites périodiques s'aplatissent sur l'équateur; nous nous sommes seulement intéressé ici à celles qui correspondent à des limites entre stabilité et instabilité, mais une infinité d'autres sont mises en évidence (N° 15); les orbites limites sont cependant les seules qui terminent des familles pour lesquelles la période de λ vaut la période de x; nous retrouvons parmi ces orbites limites, la terminaison de la famille principale (orbite 19L du travail de Lifshitz [10]), une orbite en fer à cheval du même type que deux orbites données par Störmer [13] pour $\gamma_1 = 0.97$, une orbite ovale qui termine une famille du même nom que nous avons déterminée dans notre thèse de Doctorat et dont les résultats seront publiés plus tard.

Nous tenons à remercier encore le Chanoine G. Lemaître, pour ses précieux conseils et certaines remarques que nous avons reproduites au N° 16. Nous remercions aussi le Fonds National Belge de la Recherches Scientifique qui nous a permis de faire en 1947-48 une bonne partie de ces recherches.

References

[1] L. P. Bouckaert, *Trajectoires voisines de l'équateur*, Ann. Soc. Sci. Brux., vol. 54 (1934), 174-193.

[2] L. Brillouin, *A practical method for solving Hill's equation*, Quart. of Applied Math., vol. 6 (1948), 167-178.

[3] O. Godart,. *Détermination des exposants caractéristiques des trajectoires périodiques*, Ann. Soc. Sci. Brux., vol. 58 (1938), 27-41.

[4] C. Graef et S. Kusaka, *On periodic orbits in the equatorial plane of a magnetic dipôle*, J. Math. and Phys., vol. 17 (1938), 43-54.

[5] P. Humbert, *Fonctions de Lamé et Fonctions de Mathieu* (Paris, Gauthier-Villars, 1926).

[6] G. Lemaître, *Trajectoires infiniment voisines de l'équateur*, Ann. Soc. Sci. Brux., vol. 54 (1934), 162-174.

[7] ——— *Champ magnétique et rayons cosmiques*, Ciel et Terre, Brux., vol. 59 (1943), 1-16.

[8] G. Lemaître et M. S. Vallarta, *Calcul d'une famille d'orbites asymptotiques*, Ann. Soc. Sci. Brux., vol. 56 (1936), 102-130.

[9] J. Lifshitz, *On the Fourier analysis of orbits in the equatorial plane of a magnetic dipole*, J. Math. and Phys., vol. 21 (1942), 94-116.

[10] ——— *On the stability of the principal periodic orbits in the theory of primary cosmic rays*, J. Math. and Phys., vol. 21 (1942), 284-292.

[11] F. R. Moulton, Rendiconti del Circolo Mathematico di Palermo, vol. 32 (1908), 911.

[12] ——— Monthly Notices of the R.A.S., vol. 75 (1914), 40-57.

[13] C. Störmer, *Periodische Elektronbahnen*, Zeits. fur Astroph., vol. 1 (1930), 237-274.

[14] T. Yong-Li, *Trajectoires voisines de l'équateur*, Ann. Soc. Sci. Brux., vol. 59 (1939), 301-345.

[15] M. S. Vallarta, *An outline of the theory of the allowed cone of cosmic radiation* (Univ. of Toronto Press, 1938).

[16] E. T. Whittaker, *Analytical dynamics* (Cambridge Univ. Press, 1917).

[17] E. T. Whittaker et G. N. Watson, *Modern Analysis* (Cambridge Univ. Press, 1927), chapitre XI.

Université Laval, Québec

UNION CURVES OF A HYPERSURFACE

C. E. SPRINGER

1. Introduction. A curve on an ordinary surface is a union curve[1] if its osculating plane at each point contains the line of a specified rectilinear congruence through the point. The author[2] has obtained the differential equations of union curves on a metric surface in ordinary space and has exhibited certain generalizations for union curves of known results concerning geodesic curves on a surface. It is the purpose of the present paper to develop the differential equations of the union curves of a hypersurface V_n immersed in a Riemannian manifold V_{n+1} of $n+1$ dimensions. The osculating plane to a curve on a surface is generalized to a totally geodesic surface the straight lines of which are geodesics in the space V_{n+1}. A formula is given for the union curvature vector of a curve in V_n.

2. Vector field in V_n. If y^α ($\alpha = 1, \ldots, n+1$) denote the coordinates of a point in V_{n+1}, and x^i ($i = 1, \ldots, n$) the coordinates of a point in V_n, the equations of the hypersurface V_n may be written in the form

$$y^\alpha = y^\alpha(x^1, \ldots, x^n). \tag{1}$$

For points in the V_n the functional matrix $\|\partial y^\alpha/\partial x^i\|$ is of rank n. Let the metric of V_n be denoted by $g_{ij}dx^i dx^j$ and that of V_{n+1} by $a_{\alpha\beta}dy^\alpha dy^\beta$. These metrics are assumed to be positive definite. It follows that

$$a_{\alpha\beta}y^\alpha{}_{,i}\, y^\beta{}_{,j} = g_{ij}\,, \tag{2}$$

where $y^\alpha{}_{,i}$ denotes the covariant derivative of y^α with respect to x^i. (Greek indices always have the range $1, \ldots, n+1$ and Latin indices the range $1, \ldots, n$.) If N^α denote the components of a unit vector in V_{n+1} normal to V_n, then

$$a_{\alpha\beta}y^\alpha{}_{,i}\, N^\beta = 0 \qquad (i = 1, \ldots, n), \tag{3}$$

and

$$a_{\alpha\beta}N^\alpha N^\beta = 1. \tag{4}$$

If a vector field in V_n has components U^α in the y's and components u^i in the x's, then the relation

Received July 5, 1949. Presented to the American Mathematical Society, April 30, 1949.

[1] P. Sperry, *Properties of a certain projectively defined two-parameter family of curves on a general surface*, Amer. J. of Math., vol. 40 (1928), p. 213.

[2] C. E. Springer, *Union curves and union curvature*, Bull. Amer. Math. Soc., vol. 51 (1945), pp. 686-691.

$$U^\alpha = y^\beta{}_{,i}\, u^i \tag{5}$$

must obtain. If q^α are the contravariant components in the y's of the derived vector relative to V_{n+1} of a vector of the field along a curve C in V_n, and if p^i are the contravariant components in the x's of the derived vector relative to V_n of the same vector along C, it can be shown[3] that

$$q^\alpha = \Omega_{ij}\, u^i \frac{dx^j}{ds} N^\alpha + y^\alpha{}_{,i}\, p^i, \tag{6}$$

where $\Omega_{ij}\, dx^i\, dx^j$ is the second fundamental form for V_n.

3. Totally geodesic surface in V_{n+1}. As an analogue for the osculating plane in ordinary space a totally geodesic surface in V_{n+1} is introduced. It is determined by the tangent to the curve C with equations $x^i = x^i(s)$ in V_n, s denoting arc length, and by the first curvature vector in V_{n+1} of the curve C. Let λ^α be the contravariant components in the y's of a unit vector in the direction of a curve of a congruence of curves, one curve of which passes through each point of V_n. The vector with components λ^α is, in general, not normal to V_n, and may be specified by

$$\lambda^\alpha = t^i y^\alpha{}_{,i} + rN^\alpha, \tag{7}$$

where t^i and r are parameters. Because λ^α represent a unit vector $a_{\alpha\beta}\lambda^\alpha\lambda^\beta = 1$, and it follows by use of equations (3), (4), (7) that

$$t_i t^i = 1 - r^2.$$

If the geodesic in V_{n+1} in the direction of the curve of the congruence with direction λ^α is to be a geodesic of the totally geodesic surface, then it is necessary that λ^α be a linear combination of $y^\alpha{}_{,i}\, u^i$ and q^α. Hence,

$$t^i y^\alpha{}_{,i} + rN^\alpha = v y^\alpha{}_{,i}\, u^i + w q^\alpha, \tag{8}$$

wherein v and w are to be determined, the u^i of equations (5) are now dx^i/ds, and q^α are given by

$$q^\alpha = \Omega_{ij} \frac{dx^i}{ds} \frac{dx^j}{ds} N^\alpha + y^\alpha{}_{,i}\, p^i, \tag{9}$$

and p^i are given by

$$p^i = \frac{d^2x^i}{ds^2} + \left\{ {i \atop jk} \right\} \frac{dx^j}{ds} \frac{dx^k}{ds}. \tag{10}$$

If K_n is written for $\Omega_{ij} \dfrac{dx^i}{ds} \dfrac{dx^j}{ds}$, which is the normal component of the curvature vector of the curve C in V_{n+1}, equations (8) take the form

[3] C. E. Weatherburn, *Riemannian Geometry and the Tensor Calculus* (Cambridge University Press, 1938).

$$t^i y^\alpha{}_{,i} + rN^\alpha = vy^\alpha{}_{,i}\frac{dx^i}{ds} + w(\mathrm{K}_n N^\alpha + y^\alpha{}_{,i}\, p^i). \tag{11}$$

Multiplication of equations (11) by $a_{\alpha\beta}y^\beta{}_{,j}$, summation with respect to α, and use of equations (2), (3) yield the n equations

$$g_{ij}t^i = vg_{ij}\frac{dx^i}{ds} + wg_{ij}p^i. \tag{12}$$

If equations (11) are multiplied by $a_{\alpha\beta}N^\alpha$, summation on α and use of (4) give the relation

$$r = w\mathrm{K}_n. \tag{13}$$

The solution of (12) for v is effected by multiplying by $\frac{dx^j}{ds}$ and summing on j. Because $g_{ij}p^i\frac{dx^j}{ds} = 0$, it follows that

$$v = g_{ij}t^i\frac{dx^j}{ds}. \tag{14}$$

Therefore, on using the values of v and w from (13) and (14), the n equations (12) take the form

$$g_{ij}t^i = g_{ij}\frac{dx^i}{ds}g_{lm}t^l\frac{dx^m}{ds} + \frac{r}{\mathrm{K}_n}g_{ij}p^i. \tag{15}$$

Multiplication of equations (15) by g^{jk}, summation on j, and the replacement of t^k/r by l^k lead to

$$p^k - \mathrm{K}_n\left(l^k - g_{im}l^i\frac{dx^m}{ds}\frac{dx^k}{ds}\right) = 0 \qquad (k = 1, \ldots, n), \tag{16}$$

wherein p^k are given by equations (10).

4. Union curves in V_n. For a congruence specified by the parameters l^k, the solutions of the n equations (16) determine the union curves in V_n relative to that congruence. The parameter r can not vanish under the assumption that the direction λ^α is not in the V_n. The left members of equations (16) may be denoted by η^k, which we shall call the contravariant components of the union curvature vector in V_{n+1}. A union curve of V_n with respect to a congruence determined by the parameters l^k may therefore be defined as a curve along which the union curvature vector is a null vector.

By use of (10) and the fact that $g_{ij}dx^idx^j = ds^2$, equations (16) can be written in the form

$$\eta^k \equiv p^k - \mathrm{K}_n\nu^k = 0, \tag{17}$$

where the vector ν^k is defined by

$$\nu^k = g_{ij}\frac{dx^i}{ds}\left(l^k\frac{dx^j}{ds} - l^j\frac{dx^k}{ds}\right).$$

From equations (17) it follows that if the curve C is an asymptotic curve in V_n, in which case $K_n = 0$ along the curve, then for a union curve ($\eta^k = 0$), $p^k = 0$ and the curve is a geodesic. Hence, if a union curve is an asymptotic curve, it is a geodesic. Furthermore, if a union curve is a geodesic, then it is either an asymptotic curve or the vector of components ν^k is a null vector.

The magnitude K_U of the vector η^k is given by $K_U^2 = g_{ij}\eta^i\eta^j$. From equations (7) it is seen that the angle ϕ between the vectors λ^a and N^a in V_{n+1} is given by $\cos\phi = r$, and because $t^k/r = l^k$ and $t_it^i = 1 - r^2$, it follows that $g_{ij}l^il^j = \tan^2\phi$. The angle α between the vector l^k and the tangent vector to C is given by $\cos\alpha = g_{ik}l^i\dfrac{dx^k}{ds}$. In terms of ϕ and α, the magnitude K_U of the union curvature vector can be shown to be given by

$$K_U = K_g - K_n \tan\phi \sin\alpha,$$

where K_g is the geodesic curvature of the curve C in V_n. It is to be observed that if $\phi = 0$, the union curve is a geodesic.

University of Oklahoma

INCIDENCE RELATIONS IN MULTICOHERENT SPACES II

A. H. STONE

Introduction. One standard method of studying the incidences of a system of sets $A_1, A_2, \ldots, A_n$ is to consider the nerve $\mathfrak{N}$ of the system. However, this gives no direct information as to the numbers of components of the various intersections of the sets—information which would be desirable in several geometrical problems. The object of the present paper is to modify the definition of the nerve so that these numbers of components can be taken into account, and to study this *modified nerve* $\mathfrak{M}$ for systems of sets in a connected, locally connected, normal T_1 space S of a given degree of multicoherence[1] $r(S)$. The principal result (Theorem 6, 6.4) is a refinement of a theorem of Eilenberg [4, p. 107], and asserts that, if $\bigcup A_i = S$, then under suitable hypotheses we have

$$r(\mathfrak{N}) \leqslant r(\mathfrak{M}) \leqslant r(S). \tag{1}$$

This theorem has several geometrical applications, but we shall have to leave these for subsequent treatment.

The proof proceeds as follows. After the necessary definitions (§1), we show (§2) that the modified nerve $\mathfrak{M}$ is conveniently related to the family of (continuous) mappings of S in the unit circle S^1. Next it is shown (§§3-5) that the *analytic degree of multicoherence*[2] $\rho(S)$ is equal to $r(S)$ even at the present generality; the proof, which makes frequent use of modified nerves, depends essentially on first obtaining (1) for the case in which $\mathfrak{M}$ and $\mathfrak{N}$ are 1-dimensional. The *analytic* technique of Borsuk and Eilenberg is then applied to deduce (1) in full generality, and to yield a few related results.

Though it will be clear that much of the work does not require the assumption of local connectedness, we shall use S throughout the paper to denote a non-empty, connected, locally connected, normal T_1 space. For notations in general we refer to [9] and [10].

1. The modified nerve

1.1. *Definitions, etc.* Let $A_1, A_2, \ldots, A_n$ be n given subsets of S. For each non-empty subset $J = \{i_1, i_2, \ldots, i_r\}$ of the set I of all integers from 1

Received September 20, 1949.

[1]Here $r(S) = \sup b_0(A \cap B)$, where A and B are closed connected sets such that $A \cup B = S$; the definition of b_0 is given below (footnote 4). For the fundamental properties of $r(S)$, see [3, 4, 12] in the bibliography at the end of the paper; for notations in general, see [9, 10]. In [10] the space S was assumed in addition to be completely normal; but as indicated in [10, 6.6(3)], this extra assumption is not needed for the results which will be quoted here.

[2]This notation follows [12, p. 229].

to n, we shall write A_J as an abbreviation for $A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r}$. By a ***decomposition system*** (abbreviated to d.s.) $\mathfrak{D} = \{A_J^\alpha\}$ of the system $A_1, A_2, \ldots, A_n$, we shall mean a decomposition of each A_J into a finite number[3] (possibly zero) of pairwise separated sets A_J^α with $\alpha = 1, 2, \ldots, \alpha(J)$ (so that, for each fixed J, we have $\bigcup A_J^\alpha = A_J$, $A_J^\alpha \cap A_J^\beta = 0$ if $\alpha \neq \beta$, and A_J^α is both open and closed relative to A_J), in such a way that the following "consistency" criterion is satisfied:

(1) Given α, J and J' such that $J' \subset J$, there exists α' such that $A_{J'}^{\alpha'} \supset A_J^\alpha$. (It follows that α' is unique, unless $A_J^\alpha = 0$.)

The sets $A_1, A_2, \ldots, A_n$ always have a *trivial* d.s. in which every $\alpha(J) = 1$ and $A_J^1 = A_J$. If further $A_1, A_2, \ldots, A_n$ satisfy[4] $b_0(A_J) < \infty$ for every J—or, as we shall say, if they are *of finite incidence*—they have a *natural* d.s., defined by taking the sets A_J^α to be the components of A_J. We shall be mainly interested in natural d.s.'s, though more general ones will sometimes have to be taken into account.

1.2. Corresponding to every d.s. $\mathfrak{D}$ of $A_1, A_2, \ldots, A_n$, we construct a complex $\mathfrak{M}(\mathfrak{D})$, the modified nerve of the decomposition, as follows. To each non-empty $A_{(j)}^\alpha$ we assign a vertex $a_{(j)}^\alpha$ of $\mathfrak{M}(\mathfrak{D})$ $(1 \leqslant j \leqslant n)$, and generally to each *non-empty* A_J^α we assign an open simplex a_J^α of $\mathfrak{M}(\mathfrak{D})$ having as vertices those points $a_{(j)}^{\alpha'}$ for which $j \in J$ and $A_J^\alpha \subset A_{(j)}^{\alpha'}$ (in accordance with (1) above). The *faces* of a_J^α are defined to be those simplexes $a_{J'}^{\alpha'}$ for which $J' \subset J$ and $A_{J'}^{\alpha'} \supset A_J^\alpha$; thus, for given α, J and J', there is exactly one face $a_{J'}^{\alpha'}$. With the obvious definition of incidence numbers, $\mathfrak{M}(\mathfrak{D})$ is a complex [6, p. 89] but not in general a simplicial complex [6, p. 92] (since several distinct simplexes may have identical vertices), though it becomes one on barycentric subdivision [8, p, 50]. We shall suppose $\mathfrak{M}(\mathfrak{D})$ to be realized geometrically ,and shall use $\mathfrak{M}(\mathfrak{D})$ to denote also the resulting (curved) polytope.

For the trivial d.s., $\mathfrak{M}(\mathfrak{D})$ reduces to the usual *nerve*, $\mathfrak{N}$ of $A_1, A_2, \ldots, A_n$. If the sets A_j have finite incidence and $\mathfrak{D}$ is the natural d.s., we shall write $\mathfrak{M}(\mathfrak{D})$ simply as $\mathfrak{M}$, and refer to $\mathfrak{M}$ as "the" modified nerve[5] of $A_1, A_2, \ldots, A_n$.

1.3. THEOREM 1. *Let $\mathfrak{N}$ be the nerve and $\mathfrak{M}$ the modified nerve of a system of connected sets $A_1, A_2, \ldots, A_n$ of finite incidence, and suppose that $A_j - A_k$ and $A_k - A_j$ are always separated*[6] $(1 \leqslant j, k \leqslant n)$. *Then $b_0(\mathfrak{M}) = b_0(\mathfrak{N}) = b_0(\bigcup A_j)$; and if $\bigcup A_j$, and therefore also $\mathfrak{M}$ and $\mathfrak{N}$, are connected, we have $r(\mathfrak{M}) \geqslant r(\mathfrak{N})$.*

Proof. We omit the easy argument showing that $b_0(\mathfrak{M}) = b_0(\bigcup A_j) = b_0(\mathfrak{N})$.

[3] It would be easy to extend these considerations to suitable infinite decompositions; cf. 5.3 below.

[4] Following [3], $b_0(X) + 1$ = number of components of X, if this number is finite, and $b_0(X) = \infty$ otherwise; in particular, $b_0(O) = -1$.

[5] Though $\mathfrak{M}$ consists, roughly, of $\mathfrak{N}$ with repeated cells, $\mathfrak{M}$ need not contain any subcomplex isomorphic with $\mathfrak{N}$.

[6] This condition (introduced in [11]) will always be satisfied if the sets A_j are all open, or all closed, relative to their union.

To prove $r(\mathfrak{M}) \geqslant r(\mathfrak{N})$, let the vertices of $\mathfrak{M}$ (as in 1.2) be $a_1^1, a_2^1, \dots, a_n^1$, and let those of $\mathfrak{N}$ be $a_1, a_2, \dots, a_n$, a_j^1 and a_j both corresponding to the connected set A_j. There exists an obvious simplicial mapping f of $\mathfrak{M}$ onto $\mathfrak{N}$ such that $f(a_j^1) = a_j$, and it is easy to see that any closed edge-path in $\mathfrak{N}$ is the image under f of at least one closed edge-path in $\mathfrak{M}$. Thus f induces a homomorphism of $\pi_1(\mathfrak{M})$ onto $\pi_1(\mathfrak{N})$, π_1 denoting the fundamental group. By a theorem of Eilenberg [4, p. 110] there is a homomorphism of $\pi_1(\mathfrak{N})$, and thus also of $\pi_1(\mathfrak{M})$, onto the free (non-abelian) group with $r(\mathfrak{N})$ generators; and hence [4, p. 110] $r(\mathfrak{M}) \geqslant r(\mathfrak{N})$.

2. Mappings in S^1

2.1. In what follows, f, g, etc. will denote (continuous) mappings of some normal space X (usually a subset of S) in the space S^1 of complex numbers z with $|z| = 1$; and ϕ, ψ, etc. will similarly denote continuous real-valued functions on X. To save notation, we shall usually not distinguish between a mapping $f : X \to S^1$ and the "partial mapping" $f \mid X'$ (f restricted to X') where $X' \subset X$. For the convenience of the reader, we repeat the following definitions (cf. [2], [3], [12, ch. 11]).

The *product* fg is defined by $fg(x) = f(x)g(x)$, the multiplication on the right being that of ordinary complex numbers; and the powers f^q ($q = 0, \pm 1, \pm 2, \dots$) are defined similarly. If there exists ϕ such that $f(x) = g(x) \exp (i\phi(x))$ for all $x \in X$, we write $f \sim g$ on X; in particular, if $f(x) = \exp (i\phi(x))$ we write $f \sim 1$ on X. Mappings $f_1, f_2, \dots, f_n$ are said to be (linearly) *dependent* on X if integers $q_1, q_2, \dots, q_n$ exist, positive or negative but not all zero, such that $f_1^{q_1} f_2^{q_2} \dots f_n^{q_n} \sim 1$ on X; otherwise they are *independent* on X. If $X = S$, the qualifying phrases "on X" will generally be omitted.

Given n sets $A_1, A_2, \dots, A_n$, the greatest number of mappings f of $X = \bigcup A_j$ in S^1 which satisfy

$$(1) \qquad f \sim 1 \text{ on } A_j, \qquad 1 \leqslant j \leqslant n$$

and which are independent on X (or ∞ if there is no such greatest number) is written $p(A_1, A_2, \dots, A_n)$.

Finally, the supremum of $p(F_1, F_2)$ as F_1, F_2 range over all pairs of closed sets (not necessarily connected) such that $F_1 \cup F_2 = S$, is denoted by $\rho(S)$. It is known ([3, p. 172], [4, p. 113]) that $\rho(S) = r(S)$, provided that S is a Peano space or infinite polytope; we shall later be able to remove this proviso.

2.2. Many of the arguments and results in [2], [3] (in which the space X is assumed to be metric) apply here also with, at most, trivial changes. In particular:

(1) If f maps $A \cup B$ in S^1, where the sets $A - B$ and $B - A$ are separated and $A \cap B$ is connected, and if $f \sim 1$ on A and $f \sim 1$ on B, then $f \sim 1$ on $A \cup B$ [2, p. 64, (5)].

(2) If f maps X in S^1, where X is normal, and if A is a (relatively) closed

subset of X on which $f \sim 1$, there exists a relatively open subset U of X such that $U \supset A$ and $f \sim 1$ on U([2, p. 65 (6)]; here the proof needs modification, and uses the fact that the real line is an AR [6, p. 28]).

(3) If f, g both map X in S^1 and $|f(x) - g(x)| < 1$ for each $x \in X$, then $f \sim g$ on X [3, p. 156, (2)].

(4) If f maps a closed simplex E in S^1, then $f \sim 1$ on E.

2.3. There is a close connection between modified nerves and mappings in S^1, as is shown by:

THEOREM 2. *Let $\mathfrak{M}$ be the modified nerve of a system of closed sets $A_1, A_2, \ldots, A_n$ of finite incidence. Then*[7] $b_1(\mathfrak{M}) = p(A_1, A_2, \ldots, A_n)$.

We prove (and shall need) a little more than this:

(1) If $A_1, A_2, \ldots, A_n$ are of finite incidence and such that $A_j - A_k$ and $A_k - A_j$ are always separated (but are not necessarily closed), then $b_1(\mathfrak{M}) \geqslant p(A_1, A_2, \ldots, A_n)$.

(2) If $A_1, A_2, \ldots, A_n$ are closed (but not necessarily of finite incidence), and if $\mathfrak{M} = \mathfrak{M}(\mathfrak{D})$ is the modified nerve corresponding to a d.s. $\mathfrak{D}$ of $A_1, A_2, \ldots, A_n$, then $b_1(\mathfrak{M}) \leqslant p(A_1, A_2, \ldots, A_n)$.

2.4. *Proof of* (1). First, to each mapping f of $\bigcup A_j$ in S^1 such that $f \sim 1$ on each A_j, we can assign a 1-cocycle class on $\mathfrak{M}$, as follows: We have $f(x) = \exp(i\phi_j(x))$ (say) for $x \in A_j$. For each 1-cell a_{jk}^α of $\mathfrak{M}$ (oriented from j to k), we pick $y \in A_{(j,k)}^\alpha$, and define $n_{jk}^\alpha = \{\phi_j(y) - \phi_k(y)\}/2\pi$; this number is an integer independent of the choice of y (because $A_{(j,k)}^\alpha$ is connected). It is easily verified that the 1-chain $c(f) = \sum n_{jk}^\alpha a_{jk}^\alpha$ is a cocycle, and that different choices of functions ϕ_j give rise to cocycles $c(f)$ differing only by coboundaries.

Now let μ such mappings f_λ $(1 \leqslant \lambda \leqslant \mu)$ be given, and suppose $\mu > b_1(\mathfrak{M})$. There exist integers $p_1, p_2, \ldots, p_\mu$, not all zero, such that $\sum p_\lambda c(f_\lambda) \sim 0$. Define $F = f_1^{p_1} f_2^{p_2} \ldots f_n^{p_n}$; thus we have $F \sim 1$ on each A_j, say $F = \exp(i\Phi_j)$ on A_j. Again, it readily follows that

$$c(F) = \sum N_{jk}^\alpha a_{jk}^\alpha,$$

say $\sim \sum p_\lambda c(f_\lambda) \sim 0$. Hence there exists a 0-cochain $\sum q_j^\beta a_j^\beta$ such that $N_{jk}^\alpha = q_j^\beta - q_k^\gamma$, where a_j^β, a_k^γ are the end-points of a_{jk}^α. Define a real-valued function Ψ on $\bigcup A_j$ by: $\Psi(x) = \Phi_j(x) - 2\pi q_j^\beta$ whenever $x \in A_j$. This definition is single-valued (and therefore continuous), since if $x \in A_j^\beta \cap A_k^\gamma$ we have $x \in A_{jk}^\alpha$ for some α, and then

$$(\Phi_j(x) - 2\pi q_j^\beta) - (\Phi_k(x) - 2\pi q_k^\gamma) = 2\pi(N_{jk}^\alpha - q_j^\beta + q_k^\gamma) = 0.$$

Since clearly $F = \exp(i\Psi)$ on $\bigcup A_j$, the mappings f_λ are not independent on $\bigcup A_j$ if $\mu > b_1(\mathfrak{M})$, and consequently $p(A_1, A_2, \ldots, A_n) \leqslant b_1(\mathfrak{M})$.

[7] Generalizing [2, p. 96]. Here b_1 denotes the 1-dimensional Betti number with (say) rational coefficients.

2.5. *Proof of* (2). Now let c be a given 1-cocycle on $\mathfrak{M}$, its multiplicity on the oriented 1-cell a_{jk}^{α} being the integer m_{jk}^{α} say $(= -m_{kj}^{\alpha})$. We shall define, by recursion, real-valued continuous functions Ψ_k on $A_k \cap (A_1 \cup \ldots \cup A_{k-1})$ and ϕ_k on A_k, where $k = 1, 2, \ldots, n$, setting $\phi_1 \equiv 0$ on A_1, $\psi_2 = -2\pi m_{12}^{\alpha} + \phi_1$ on A_{12}^{α} $(\alpha = 1, 2, \ldots, a(12))$, $\phi_2 =$ an extension of ψ_2 to A_2, and generally

$$\psi_k = -2\pi m_{jk}^{\alpha} + \phi_j \text{ on } A_{jk}^{\alpha} \qquad (1 \leqslant j < k, 1 \leqslant \alpha \leqslant a(jk)),$$

and $\phi_k =$ an extension of ψ_k to A_k. To justify this definition, we must first show that the definition of ψ_k is consistent, i.e., that if $h < j < k$ and $x \in A_h \cap A_j \cap A_k$, say $x \in A_{hjk}^{\delta} \subset A_{jk}^{\alpha} \cap A_{hk}^{\beta} \cap A_{jk}^{\gamma}$, then

$$-2\pi m_{jk}^{\alpha} + \phi_j(x) = -2\pi m_{hk}^{\beta} + \phi_h(x).$$

This follows from the fact that $m_{jk}^{\alpha} + m_{kh}^{\beta} + m_{hj}^{\gamma} = 0$, c being a cocycle. Since ψ_k is thus a well-defined continuous function on the closed subset $A_k \cap (A_1 \cup \ldots \cup A_{k-1})$ of the normal space A_k, the extension ϕ_k exists [6, p. 28].

It follows that, whenever $x \in A_j \cap A_k$, we have $\exp(i\phi_j(x)) = \exp(i\phi_k(x))$; consequently the mapping f defined by

$$f = \exp(i\phi_j) \text{ on } A_j, \qquad 1 \leqslant j \leqslant n$$

is single-valued and continuous on $\bigcup A_j$. Further, even though the sets A_{jk}^{α} need not now be connected, we have $\phi_j(y) - \phi_k(y) = 2\pi m_{jk}^{\alpha}$ whenever $y \in A_{jk}^{\alpha}$, so that a cocycle $c(f)$ can still be associated with f as in 2.4 above, and is evidently simply c.

Now let $b_1(\mathfrak{M}) = \mu$, and choose μ 1-cocycles c_λ, $1 \leqslant \lambda \leqslant \mu$, linearly independent modulo cohomology in $\mathfrak{M}$. Corresponding to each c_λ, the above construction gives a mapping f_λ of $\bigcup A_j$ in S^1 such that

$$\text{(i)} \quad f_\lambda \sim 1 \text{ on each } A_j, \qquad \text{(ii)} \quad c(f_\lambda) = c_\lambda.$$

We have only to show that theses mappings f_λ are independent on $\bigcup A_j$. But if say $F \equiv f_1^{q_1} f_2^{q_2} \ldots f_\mu^{q_\mu} = \exp(i\Phi)$ on $\bigcup A_j$, where the q_j's are integers and Φ is a continuous real-valued function, we readily see that $c(F)$ exists and $\sum q_\lambda c_\lambda \sim c(F) \sim 0$; hence $q_1 = q_2 = \ldots = 0$.

2.6. COROLLARY. *If $A_1, A_2, \ldots, A_n$ are closed sets of finite incidence, no three of which have a common point, then*[8] $p(A_1, A_2, \ldots, A_n) = b_1(\mathfrak{M}) = h(A_1, A_2, \ldots, A_n)$.

For the definition of $h(A_1, A_2, \ldots, A_n)$ here reduces to

$$b_0(\textstyle\bigcup A_j) + \sum_{j<k} (b_0(A_j \cap A_k) + 1) - n + 1 - \sum b_0(A_j).$$

Now $\mathfrak{M}$ is a linear graph having $b_0(\bigcup A_j) + 1$ components, $\sum(b_0(A_j \cap A_k) + 1)$ edges, and $\sum b_0(A_j) + n$ vertices; hence $h(A_1, A_2, \ldots A_n) = b_1(\mathfrak{M})$, by the Euler-Poincaré formula.

[8]By definition [9, p. 441], $h(A_1, \ldots, A_n) = \sum_1^n b_0(X_r) - \sum_1^n b_0(A_j)$, where X_r is the set of all points belonging to A_j for r or more values of j.

Remark. For closed sets in general we have

$$p(A_1, A_2, \ldots, A_n) \leqslant h(A_1, A_2, \ldots, A_n).$$

This can be proved by induction over n, the case $n = 2$ being furnished by the above corollary.

3. Lemmas on linear graphs

3.1. In the next section we shall study "one-dimensional" coverings of S, whose modified nerves will be linear graphs; in preparation for this, we here collect the necessary graph-theoretic lemmas. In view of the applications, a (linear) *graph* G will here mean a finite 1-complex which may be "improper", i.e., in which two vertices may be joined by several edges (open 1-cells); but each edge is to have two distinct vertices. We denote the numbers of vertices and edges of G by $a_0(G)$, $a_1(G)$ respectively. The *order* $\nu(p, G)$ of a vertex p of G is the number of edges of G which are incident with p (have p as a vertex). A vertex p of order 1 is an *end-point* of G, and the single edge incident with p is then an *end-line*. An acyclic connected non-empty graph is a *tree*.

3.2. From the Euler-Poincaré formula, combined with the equality of b_1 and r for 1-dimensional Peano spaces [3, p. 162], we have:

(1) If G is a connected and non-empty graph, then

$$a_1(G) - a_0(G) + 1 = b_1(G) = r(G).$$

An elementary computation then gives:

(2) If G is a tree having exactly λ end-points and μ other vertices $q_1, q_2, \ldots, q_\mu$, then

$$\textstyle\sum_1^\mu \{\nu(q_j, G) - 2\} = \lambda - 2.$$

We note also the obvious property:

(3) If G is a connected graph having an end-point p with end-line C, then $G - C - (p)$ is connected.

3.3. Now let G be a graph having vertices $p_1, p_2, \ldots, p_m$ and edges $C_1, C_2, \ldots, C_n$, and suppose there exists a (continuous) monotone simplicial mapping ϖ of a graph H onto G. Thus $\varpi^{-1}(p_j)$ is a (closed) connected subgraph of H, $\varpi^{-1}(C_k)$ is a single (open) edge of H, and these inverse sets are pairwise disjoint, non-empty, and cover H. Suppose further that whenever C_j, C_k are distinct edges of G, the edges $\varpi^{-1}(C_j)$, $\varpi^{-1}(C_k)$ have disjoint closures (i.e., have no end-point in common). We shall then call ϖ^{-1} a *dispersion of* G, and shall also say that H is a dispersion of G. (Roughly speaking, the operation of "dispersing" G into H consists in replacing the vertices p_j of G by disjoint connected graphs $\varpi^{-1}(p_j)$, and reattaching the 1-cells of G in such a way that no two of them have a common vertex.)

3.4. In what follows, we suppose that H is a dispersion of a *connected* graph G. Since ϖ is monotone,

(1) $$H \text{ is connected};$$

and from 3.2(1) we readily obtain

(2) $$b_1(H) \geqslant b_1(G).$$

A dispersion of G will be called *minimal* if it satisfies: (a) the sub-graphs $\varpi^{-1}(p_j)$ are all trees, (b) each end-point of each $\varpi^{-1}(p_j)$ is incident with at least one (and therefore exactly one) edge $\varpi^{-1}(C_k)$. From 3.2(1) we see that:

(3) If H is a minimal dispersion of G, then

$$b_1(H) = b_1(G).$$

Further,

(4) Given a dispersion H of G, and a subgraph G^* of G, there exists a subgraph H^* of H which is a minimal dispersion of G^*.

In fact, $H_1 = \varpi^{-1}(G^*)$ is a subgraph of H which is a dispersion of G^*. Of those subgraphs of H_1 which are dispersions of G^*, let H^* be one having as few edges as possible. It is easy to see that H^* will be a minimal dispersion of G^*.

Now assume that H is a *minimal* dispersion of a connected graph G, and let the vertices of the subgraph $\varpi^{-1}(p)$ of H (p being a given vertex of G) be $q_1, q_2, \ldots, q_h$. An easy calculation, based on 3.2(2), gives $\sum_1^h\{\nu(q_j, H) - 2\} = \nu(p, G) - 2$, whence, since (with trivial exceptions) each summand is non-negative:

(5) $$\nu(q_j, H) \leqslant \nu(p, G) \qquad (1 \leqslant j \leqslant h);$$

and if for some j we have $\nu(q_j, H) = \nu(p, G)$, then

$$\nu(q_k, H) = 2 \text{ for all } k \neq j \qquad (1 \leqslant k \leqslant h).$$

We shall say that a minimal dispersion H of a connected graph G is *non-trivial* if there exists a vertex p of G for which the vertices q_j of $\varpi^{-1}(p)$ all satisfy $\nu(q_j, H) < \nu(p, G)$, and that it is *trivial* otherwise. From (5) we have:

(6) *If $G_1, G_2, \ldots$ is an infinite sequence of connected graphs such that G_{n+1} is a minimal dispersion of G_n ($n = 1, 2, \ldots$), then, for all large enough n, G_{n+1} is a trivial dispersion of G_n.*

Further, (5) shows that a trivial minimal dispersion of G is essentially a "subdivision" of G. In fact, we have:

(7) *Let $G_1, G_2, \ldots, G_n$ ($n \geqslant 2$) be connected graphs such that G_{j+1} is a trivial minimal dispersion of G_j ($1 \leqslant j \leqslant n - 1$). Then each non-zero 1-cycle of G_n contains (i.e., has non-zero multiplicity on) a sequence of edges $E_1, E_2, \ldots, E_m$, where $m = 2^{n-2} + 1$, such that*

(i) *E_j and E_{j+1} have exactly one common end-point, which is moreover of order 2 in G_n $(1 \leqslant j \leqslant m-1)$, and*[9]

(ii) $$\mathrm{Cl}(E_j) \cap \mathrm{Cl}(E_k) = 0 \text{ if } |j-k| \geqslant 2 \qquad (1 \leqslant j, k \leqslant m).$$

The proof of (7) is straightforward by induction over n, using (5).

4. One-dimensional coverings

4.1. THEOREM 3. *Let $r(S)$ be finite, and let $A_1, A_2, \ldots, A_n$ be n non-empty closed connected sets covering S, no three of which have a common point. Then the sets A_j are of finite incidence; and if $\mathfrak{M}$ is their modified nerve, we have $r(\mathfrak{M}) \leqslant r(S)$.*

We have, if $j \neq k$,

$$\mathrm{Fr}(A_j) \cap \mathrm{Fr}(A_k) \cap \mathrm{Fr}(A_j \cup A_k) \subset A_j \cap A_k \cap \mathrm{Cl}(\mathrm{Co}(A_j \cup A_k))$$
$$\subset A_j \cap A_k \cap \bigcup \{A_m | m \neq j, k\} = 0;$$

hence [10, 7.3] $b_0(A_j \cap A_k) \leqslant r(S) < \infty$. Thus the sets A_j are of finite incidence, and $\mathfrak{M}$ is defined (and is evidently a graph). In accordance with the notation of 1.1, we write A_{jk}^{α} $(1 \leqslant \alpha \leqslant \alpha(j,k))$ for the components of $A_{jk} = A_j \cap A_k$. Since

$$\mathrm{Fr}(A_j) \subset A_j \cap \bigcup \{A_k | k \neq j\} = \bigcup_{k,\alpha} A_{jk}^{\alpha},$$

a union of pairwise disjoint closed connected (non-empty) sets, there exist [10, 3.4], for each fixed j, closed connected sets $H_{jk}^{\alpha} \supset A_{jk}^{\alpha}$ such that $\bigcup_{k,\alpha} H_{jk}^{\alpha} = A_j$, no three of the sets H_{jk}^{α} have a common point, and the intersection of every two of them is contained in $A_j - \bigcup \{A_k | k \neq j\}$. (Note that $H_{jk}^{\alpha} \neq H_{kj}^{\alpha}$, though of course $A_{jk}^{\alpha} = A_{kj}^{\alpha}$.)

It readily follows that no three of *all* the sets H_{jk}^{α} can have a common point, even if j varies. Thus if we renumber the sets H_{jk}^{α}, say as $A_1(1), A_2(1), \ldots, A_{n_1}(1)$, the sets $A_j(1)$ have all the properties which were postulated for the sets A_j; hence they are also of finite incidence. Let the nerve and modified nerve of $\{A_j(1)\}$ be G_1 and H_1 respectively; both are graphs. We assert:

(1) $$G_1 \text{ is a dispersion of } \mathfrak{M}.$$

In fact, we can map G_1 on $\mathfrak{M}$ as follows. Each vertex q of G_1 corresponds to some set H_{jk}^{α}; we define $\varpi(q) = a_j$, the vertex of $\mathfrak{M}$ corresponding to A_j. Each edge of G_1 corresponds to a non-empty intersection $H_{jk}^{\alpha} \cap H_{lm}^{\beta}$. If $j = l$, we map the whole edge on a_j; if $j \neq l$, we must have $m = j$, $k = l$ and $\alpha = \beta$, and map the edge "linearly" onto the edge a_{jk}^{α} of $\mathfrak{M}$. The resulting mapping ϖ is easily seen to be continuous. Further, it is monotone, since ϖ^{-1} is clearly $1-1$ on a_{jk}^{α}, while $\varpi^{-1}(a_j)$ is precisely the nerve of the sets H_{jk}^{α} with fixed j, and is connected since A_j is connected. And it is not hard to see that if a_{jk}^{α} and a_{lm}^{β} are distinct edges of $\mathfrak{M}$, their inverse images under ϖ cannot have a common end-point. Thus (1) is established.

[9] We use the customary abbreviations Cl for *closure*, Co for *complement*, Fr for *frontier*.

Clearly also, to within isomorphism,

(2) $$G_1 \text{ is a subgraph of } H_1.$$

The whole process is now repeated, starting with the sets $A_j(1)$; and so on. We thus obtain, for each λ $(= 1, 2, \ldots)$, a covering of S by closed connected sets $A_j(\lambda)$ $(1 \leqslant j \leqslant n_\lambda)$, no three of which have a common point, having nerve G_λ and modified nerve H_λ, such that $G_{\lambda+1}$ is both a dispersion of H_λ and a subgraph of $H_{\lambda+1}$.

From 3.4(4), we obtain recursively a sequence of graphs K_λ such that $K_1 = G_1$ and K_λ is a subgraph of G_λ which is a minimal dispersion of $K_{\lambda-1}$ $(\lambda \geqslant 2)$. By 3.4(6), there exists an integer $N > 3$ such that $K_{\lambda+1}$ is a trivial minimal dispersion of K_λ whenever $\lambda \geqslant N - 2$. On applying 3.4(7) to K_{N-2}, K_{N-1}, K_N, we see that every non-zero 1-cycle of K_N contains a sequence C_1, C_2, C_3 of three edges, such that:

(i) $\bar{C}_1 \cap \bar{C}_2 =$ a single vertex p_1 of K_N,

(ii) $\bar{C}_2 \cap \bar{C}_3 =$ a single vertex p_2 of K_N,

(iii) $\nu(p_1, K_N) = 2 = \nu(p_2, K_N)$,

(iv) $\bar{C}_1 \cap \bar{C}_3 = 0$.

For short we shall call such a sequence of three edges a "triad".

The graph K_N is connected (3.4(1) and Theorem 1, 1.3); hence if it is not already a tree it contains a cycle containing a triad (C_1^1, C_2^1, C_3^1). The subgraph $K_N - C_2^1$ is clearly connected, and has C_1^1 and C_3^1 among its end-lines. Hence if $K_N - C_2^1$ is not a tree it contains a triad (C_1^2, C_2^2, C_3^2) disjoint from the first. After a finite number of steps, say r, we obtain r mutually exclusive triads (C_1^s, C_2^s, C_3^s), $1 \leqslant s \leqslant r$, in K_N, such that $K_N - \bigcup C_2^s = T$, say, is a tree having all the edges C_1^s, C_3^s among its end-lines.

From 3.2(1) we obtain

(3) $$r = r(K_N).$$

Let U_i denote the subgraph of T formed by omitting from T all the edges C_i^s and the corresponding end-points $\mathrm{Cl}(C_i^s) \cap \mathrm{Cl}(C_2^s)$ $(i = 1, 3)$. From 3.2(3), U_1 and U_3 are connected subgraphs of K_N, and thus *a fortiori* of G_N; further, $U_1 \cap U_3 \neq 0$, and we note that G_N also contains the r distinct edges C_2^s, no two of which have a common end-point, and each of which joins a vertex in $U_1 - U_3$ to a vertex in $U_3 - U_1$.

For each vertex p of $G_N - (U_1 \cup U_3)$, join p to a vertex of $U_1 \cup U_3$ by a simple edge-path $W(p)$ in G_N (this is possible since G_N is connected, by 1.3), and further choose $W(p)$ to have as few edges as possible. Define $V_i =$ union of U_i with all those paths $W(p)$ whose ends (other than p) are in U_i $(i = 1, 3)$. Clearly $V_1 - V_3 \supset U_1 - U_3$ and $V_3 - V_1 \supset U_3 - U_1$; and moreover $V_1 \cup V_3$ contains all the vertices of G_N. Now G_N is the nerve of the closed connected sets $A_j(N)$ covering S. Let $X_i =$ union of those sets $A_j(N)$ which correspond to vertices in V_i $(i = 1, 3)$. It readily follows that X_1, X_3 are closed connected sets which cover S, and hence

$$(4) \qquad b_0(X_1 \cap X_2) \leqslant r(S).$$

We may suppose the notation so chosen that A_j corresponds to a vertex in $V_1 \cap V_2$ if $1 \leqslant j \leqslant \mu$, in $V_1 - V_2$ if $\mu < j \leqslant \nu$, and in $V_2 - V_1$ if $\nu < j \leqslant n_N$. Write $D = A_1(N) \cup A_2(N) \cup \ldots \cup A_\mu(N)$. Then clearly

$$(5) \qquad X_1 \cap X_2 = D \cup \bigcup \{A_j(N) \cap A_k(N) | \mu < j \leqslant \nu < k \leqslant n_N\}.$$

The sets D, $A_j(N) \cap A_k(N)$ appearing here are closed and pairwise disjoint (for no three of the sets $A_j(N)$ have a common point). Further, $D \neq 0$ (for $V_1 \cap V_2 \neq 0$), and at least r of the sets $A_j(N) \cap A_k(N)$ are non-empty —namely those corresponding to the edges C_2^* of G_N. Hence (5) shows that $b_0(X_1 \cap X_2) \geqslant r$, and so, from (3) and (4), we have $r(K_N) \leqslant r(S)$. But 3.4(2) and 3.4(3) show that

$$r(\mathfrak{M}) \leqslant r(G_1) = r(K_1) = r(K_2) = \ldots = r(K_N),$$

and consequently $r(\mathfrak{M}) \leqslant r(S)$.

4.2. Corollary. *If $r(S) < \infty$, there exists a covering of S by a finite number of closed connected sets A_j, no three of which have a common point, such that* (i) *their nerve $\mathfrak{N}$ satisfies $r(\mathfrak{N}) = r(S)$, and* (ii) *every intersection $A_j \cap A_k$ is connected.*

4.3. We next derive, for later use, a related property of *open* sets (which need not necessarily cover S).

Lemma. *Let $A_1, A_2, \ldots, A_n$ be n non-empty closed connected sets such that* $\mathrm{Fr}(A_j) \cap \mathrm{Fr}(A_k) \cap \mathrm{Fr}(A_j \cup A_k) = 0$ *whenever $j \neq k$, and no three of which have a common point. Then*

$$b_0(\bigcup A_j) + b_0(\bigcup\{A_j \cap A_k | j \neq k\}) \leqslant r(S) + n - 2.$$

We may evidently assume $n > 1$ and $r(S) < \infty$. Write $U = \mathrm{Co}(\bigcup A_j)$ and $F_j = \mathrm{Fr}(U) \cap \mathrm{Fr}(A_j)$; thus $\bigcup F_j = \mathrm{Fr}(U)$ and the sets F_j are pairwise disjoint. By [10, 3.4], there exist closed sets H_j ($1 \leqslant j \leqslant n$) such that $H_j \supset F_j$, H_j is connected relative[10] to F_j, $\bigcup H_j = \overline{U}$, $H_j \cap F_k = 0$ if $j \neq k$, $H_j \cap H_k \subset U$ if $j \neq k$, and no three of the sets H_j have a common point.

Write $A_j \cup H_j = B_j$; thus the n sets B_j are closed, connected, and cover S, and no three of them have a common point. Hence, from Theorem 3 (4.1) and 3.2(1), the modified nerve $\mathfrak{M}$ of $B_1, B_2, \ldots, B_n$ exists and satisfies

$$(1) \qquad r(\mathfrak{M}) = a_1(\mathfrak{M}) - a_0(\mathfrak{M}) + 1 \leqslant r(S), \quad a_0(\mathfrak{M}) = n.$$

Now if $j \neq k$ we have $A_j \cap H_k = A_j \cap \overline{U} \cap H_k \subset F_j \cap H_k = 0$, and similarly $A_k \cap H_j = 0$. Thus

$$(2) \qquad B_j \cap B_k = (A_j \cap A_k) \cup (H_j \cap H_k);$$

and since $H_j \cap H_k \subset U$, the closed sets $A_j \cap A_k$ and $H_j \cap H_k$ are disjoint.

[10]For the definition and elementary properties of relative connectedness, see [9, p. 428] and [10, 3.3].

Thus the modified nerve $\mathfrak{M}_0$ of $A_1, A_2, \ldots, A_n$ exists and can be obtained from $\mathfrak{M}$ merely by deleting certain edges of $\mathfrak{M}$ (corresponding to the components of the sets $H_j \cap H_k$). Since $\mathfrak{M}$ is connected, while $b_0(\mathfrak{M}_0) = b_0(\bigcup A_j)$ (Theorem 1, 1.3), the number of edges so deleted must be at least $b_0(\bigcup A_j)$. Thus we have $b_0(\bigcup A_j) + a_1(\mathfrak{M}_0) \leqslant a_1(\mathfrak{M})$; and since the sets $A_j \cap A_k$ $(j < k)$ are pairwise disjoint, $a_1(\mathfrak{M}_0)$ = number of components of $\bigcup(A_j \cap A_k) = b_0(\bigcup(A_j \cap A_k)) - 1$. The lemma now follows from (1).

4.4. THEOREM 4. *Let U, V be open subsets of S which satisfy* $\mathrm{Fr}(U) \cap \mathrm{Fr}(V) \cap \mathrm{Fr}(U \cap V) = 0$. *Then* $h(U, V) \leqslant r(S)$ (i.e., $b_0(U \cup V) + b_0(U \cap V) \leqslant b_0(U) + b_0(V) + r(S)$).

Proof. We may assume that $r(S)$, $b_0(U)$ and $b_0(V)$ are all finite. Let U, V have components $U_1, \ldots, U_m$, $V_1, \ldots, V_n$ respectively. From [10, 7.4], each of the sets $U_j \cap V_k$ has only a finite number of components, say W_{jk}^α $(1 \leqslant \alpha \leqslant \alpha(jk))$. Pick points $x_j \in U_j$, $y_k \in V_k$, $z_{jk}^\alpha \in W_{jk}^\alpha$. Since U_j is open and connected, there exists a closed connected set joining x_j and z_{jk}^α in U_j; let the union of these closed connected sets, as k and α vary, be denoted by A_j. Similarly we construct a closed connected set $B_k \subset V_k$ containing all the points z_{jk}^α(for each fixed k). Write $\bigcup A_j = A$, $\bigcup B_k = B$. Then $\mathrm{Co}(A)$ and $\mathrm{Co}(B)$ are open sets containing $\mathrm{Co}(U)$ and $\mathrm{Co}(V)$ respectively; and [10, 6.3] gives the existence of open sets C, D such that $\mathrm{Co}(A) \supset C \supset \mathrm{Co}(U)$, $\mathrm{Co}(B) \supset D \supset \mathrm{Co}(V)$, and $\mathrm{Fr}(C) \cap \mathrm{Fr}(D) = 0$. Thus $A \subset \mathrm{Co}(C) \subset U$, which shows that each component A_j of A is contained in a component C_j (say) of $\mathrm{Co}(C)$, and that $C_j \subset U_j$. Similarly we obtain n distinct components D_k of $\mathrm{Co}(D)$ such that $B_k \subset D_k \subset V_k$. We have $\mathrm{Fr}(C_j) \cap \mathrm{Fr}(D_k) \subset \mathrm{Fr}(C) \cap \mathrm{Fr}(D) = 0$, so that the sets $C_1, \ldots, C_m, D_1, \ldots, D_n$ satisfy the hypotheses of the lemma (4.3), and therefore

$$(1) \qquad b_0(\bigcup C_j \cup \bigcup D_k) + b_0(\bigcup(C_j \cap D_k)) \leqslant r(S) + m + n - 2.$$

Now the different sets $C_j \cap D_k$ are pairwise disjoint, and, since $z_{jk}^\alpha \in C_j \cap D_k \subset U_j \cap V_k$, each set $C_j \cap D_k$ has at least as many components as $U_j \cap V_k$. Thus

$$b_0(\bigcup(C_j \cap D_k)) \geqslant b_0(U \cap V).$$

Similarly $b_0(\bigcup C_j \cup \bigcup D_k) \geqslant b_0(U \cup V)$; and the theorem now follows from (1).

4.5. *Remark.* A similar argument will apply to any finite number of open sets, no three of which have a common point, and every two of which satisfy the frontier relation of Theorem 4. Further, if S is completely normal, the "approximation" method [10, 6.5] can be carried a step farther [10, 7.5] to yield the following theorem:

THEOREM 4a. *If S is completely normal, and $E_1, E_2, \ldots, E_n$ are n sets, no three of which have a common point, and every two of which satisfy* (i) $E_j - E_k$ *and* $E_k - E_j$ *are separated,* (ii) $E_j \cap E_k$ *and* $\mathrm{Co}(E_j \cup E_k)$ *are separated* $(j \neq k)$, *then*

$$\sum b_0(E_j) + n - 2 \leqslant b_0(\bigcup E_j) + b_0(\bigcup\{E_j \cap E_k | j \neq k\}) \leqslant \sum b_0(E_j) + r(S) + n - 2.$$

5. The analytic definition of $r(S)$

5.1. The number $\rho(S)$, defined (2.1) in terms of mappings of S in S^1, is known to equal $r(S)$ for e.g. Peano spaces. We shall now show that this equality holds for all connected, locally connected, normal T_1 spaces, without any requirements of compactness or completeness.

THEOREM 5. $\rho(S) = r(S)$.

5.2. *Proof.* It is easy to see that

(1) $$r(S) \leqslant \rho(S).$$

In fact, let A_1, A_2 be closed connected sets which cover S, and suppose $b_0(A_1 \cap A_2) \geqslant n$. We can write $A_1 \cap A_2$ as a union of $n + 1$ disjoint closed non-empty sets A_{12}^*; and this defines a d.s. of A_1, A_2 for which the corresponding modified nerve $\mathfrak{M}$ has 2 vertices and $n + 1$ edges, so that $n = b_1(\mathfrak{M})$. But (2.3 (2)) $b_1(\mathfrak{M}) \leqslant p(A_1, A_2) \leqslant \rho(S)$; thus $n \leqslant \rho(S)$, and (1) follows.

5.3. Now suppose

(2) $$r(S) < \rho(S);$$

we shall derive a contradiction. From (2), $r(S) = n$ say $< \infty$, and there exist closed (but not necessarily connected) sets F_1, F_2 and $n + 1$ independent (continuous) mappings f_j of S in S^1 $(1 \leqslant j \leqslant n + 1)$ such that $f_j \sim 1$ on each of F_1, F_2. There exist (2.2 (2)) open sets $A \supset F_1$, $B \supset F_2$, and continuous real-valued functions ϕ_j, ψ_j, such that

(3) $$f_j = \exp(i\phi_j) \text{ on } A, \text{ and } f_j = \exp(i\psi_j) \text{ on } B \qquad (1 \leqslant j \leqslant n + 1).$$

Let A, B have components $\{A_\lambda\}$, $\{B_\mu\}$, repectively. Each of these components is open; further, we have

$$\mathrm{Fr}(A_\lambda) \cap \mathrm{Fr}(B_\mu) \subset \mathrm{Fr}(A) \cap \mathrm{Fr}(B) \subset \mathrm{Co}(A) \cap \mathrm{Co}(B) = 0.$$

Hence for any finite unions $\mathfrak{A} = A_{\lambda_1} \cup A_{\lambda_2} \cup \ldots \cup A_{\lambda_h}$ and $\mathfrak{B} = B_{\mu_1} \cup B_{\mu_2} \cup \ldots \cup B_{\mu_k}$ we have $\mathrm{Fr}(\mathfrak{A}) \cap \mathrm{Fr}(\mathfrak{B}) = 0$ and therefore (Theorem 4, 4.4)

(4) $$h(\mathfrak{A}, \mathfrak{B}) \leqslant n.$$

In particular,

(5) $$b_0(A_\lambda \cap B_\mu) \leqslant n.$$

Now form a "graph" $\mathfrak{M}$ (which however will be infinite, in general) by taking vertices a_λ, b_μ corresponding to the sets A_λ, B_μ, and joining a_λ to b_μ by as many edges as $A_\lambda \cap B_\mu$ has components. (Thus $\mathfrak{M}$ is the "modified nerve" of A and B except that it is formed with respect to an infinite decompo-

sition, in general.) From (4) and 3.2 (1) we have $b_1(G) \leqslant n$ whenever G is a subgraph of $\mathfrak{M}$ generated by a finite number of vertices of $\mathfrak{M}$, and hence also whenever G is any finite subgraph of $\mathfrak{M}$. Thus there is a finite subgraph G_1 of $\mathfrak{M}$ for which $b_1(G_1)$ is as large as possible; say $b_1(G_1) = N$, where $N \leqslant n$. Next, since $\mathfrak{M}$ is connected (for S is), there exists a connected finite subgraph G_2 of $\mathfrak{M}$ containing G_1 (obtained by adding to G_1 a finite number of edge-paths connecting the vertices of G_1 in $\mathfrak{M}$). Let $a_{\lambda_1}, \ldots, a_{\lambda_k}, b_{\mu_1}, \ldots, b_{\mu_k}$ be the vertices of G_2, and let G_3 be the subgraph of $\mathfrak{M}$ which they generate. Thus G_3 is a connected finite graph, and since $G_3 \supset G_2 \supset G_1$ we have $b_1(G_3) \geqslant b_1(G_1) \geqslant N$, and therefore $b_1(G_3) = N$. Write $\mathfrak{A} = A_{\lambda_1} \cup \ldots \cup A_{\lambda_k}$, $\mathfrak{B} = B_{\mu_1} \cup \ldots \cup B_{\mu_k}$; then clearly $\mathfrak{A}$ and $\mathfrak{B}$ are of finite incidence, and their modified nerve is G_3.

We shall next assign a "rank" to each vertex of $\mathfrak{M}$, as follows. Let $p(= a_\lambda$ or $b_\mu)$ be a given vertex of $\mathfrak{M}$. If $p \in G_3$, its rank is zero. If p non $\in G_3$, join p to G_3 by a finite edge-path $W(p)$ in $\mathfrak{M}$ such that $W(p)$ contains no edge in G_3 (e.g., take $W(p)$ to be as short as possible). We assert that $W(p)$ is now unique. In fact, if $W'(p)$ were a different edge-path satisfying these requirements, the subgraph $G_3 \cup W(p) \cup W'(p)$ would (as is easy to see) contain a closed path not lying entirely in G_3, so that $b_1(G_3 \cup W(p) \cup W'(p)) > b_1(G_3) = N$, contradicting the definition of N. The rank of p is now defined to be the number of edges in $W(p)$.

The "rank" of a component A_λ or B_μ of A or B is defined to be the rank of the corresponding vertex of $\mathfrak{M}$, and we write C_ν = union of all sets (A_λ or B_μ) of rank $\leqslant \nu$ ($\nu = 0, 1, 2, \ldots$). Thus C_ν is open, $\mathfrak{A} \cup \mathfrak{B} = C_0 \subset C_1 \subset C_2 \subset \ldots$, and $\bigcup C_\nu = S$. Further, the construction shows that the sets of fixed rank $\nu > 0$ are pairwise disjoint, while each set of rank $\nu > 0$ intersects one and only one set of rank $\nu - 1$, and this intersection is always connected.

We have $n + 1 > N = b_1(G_3) \geqslant p(\mathfrak{A}, \mathfrak{B})$, from 2.3 (1); hence, in view of (3) above, there must exist integers $q_1, q_2, \ldots, q_{n+1}$, not all zero, and a continuous real-valued function θ on $\mathfrak{A} \cup \mathfrak{B}$, such that

$$F = f_1^{q_1} f_2^{q_2} \ldots f_{n+1}^{q_{n+1}} = \exp(i\theta) \text{ on } \mathfrak{A} \cup \mathfrak{B}. \tag{6}$$

Using (3), we define

$$\Phi = \sum q_j \phi_j \text{ on } A, \Psi = \sum q_j \psi_j \text{ on } B;$$

thus $F = \exp(i\Phi)$ on A, and $F = \exp(i\Psi)$ on B.

We now extend θ to a continuous function Θ, defined for all $x \in S$, and such that $F = \exp(i\Theta)$, as follows. On C_0, define $\Theta = \theta$. Now suppose Θ has been defined with the desired properties on C_ν. If A_λ is a set of rank $\nu + 1$, it intersects a unique set of rank ν, necessarily of the form B_μ, and $A_\lambda \cap C_\nu = A_\lambda \cap B_\mu$ which is connected. Hence on $A_\lambda \cap C_\nu$ we have $\Theta = \Phi - 2\pi m_\lambda$ where m_λ is a (constant) integer; and we define $\Theta = \Phi - 2\pi m_\lambda$ on A_λ. Similarly Θ is defined on each B_μ of rank $\nu + 1$ (using the function Ψ). Since the sets of rank $\nu + 1$ are pairwise disjoint open sets, Θ is single valued and con-

tinuous on $C_{\nu+1}$, and clearly $\exp(i\Theta) = F$ on $C_{\nu+1}$. This process defines Θ with the above properties on all of S; but this contradicts the independence of the mappings f_j, and the proof is complete.

6. Finite coverings in general

6.1. LEMMA 1. *Given a d.s.* $\{A_J{}^\alpha\}$ *of* n *closed sets* $A_1, A_2, \ldots, A_n$, *and given open sets* $U(J, \alpha) \supset A_J{}^\alpha$, *there exist open* F_σ *sets* $B_1, B_2, \ldots, B_n$ *and a d.s.* $\{B_J{}^\alpha\}$ *of* $B_1, B_2, \ldots, B_n$, *such that* (i) $A_J{}^\alpha \subset B_J{}^\alpha \subset \mathrm{Cl}(B_J{}^\alpha) \subset U(J, \alpha)$, (ii) $B_J{}^\alpha$ *is connected*[10] *relative to* $A_J{}^\alpha$, (iii) $\mathrm{Cl}(B_J{}^\alpha) \cap \mathrm{Cl}(B_{J'}{}^{\alpha'}) = 0$ *whenever* $A_J{}^\alpha \cap A_{J'}{}^{\alpha'} = 0$, (iv) $B_J{}^\alpha \subset B_{J'}{}^{\alpha'}$ *whenever* $A_J{}^\alpha \subset A_{J'}{}^{\alpha'}$, *and* (v) $\mathrm{Cl}(B_J) = \bigcap \{\mathrm{Cl}(B_j) | j \in J\}$.

Remark. It follows that $\{\mathrm{Cl}(B_J{}^\alpha)\}$ will be a d.s. of the sets $\overline{B}_j$, and that if the sets A_j have finite incidences then so do the sets B_j and the sets $\overline{B}_j$, and all three systems of sets have then the same modified nerve.

Proof. Let k be the greatest number of different suffixes j, $1 \leqslant j \leqslant n$, for which the intersection of the corresponding sets A_j is not empty. The proof will go by induction over k (n remaining fixed). If $k = 1$, the result follows easily from the following two well-known properties:

(1) Given $F \subset U$, where F is closed and U open, there exists an open F_σ set V such that $F \subset V \subset \overline{V} \subset U$.

(2) If E is an open F_σ set, so is every union of components of E.

Now assume the lemma holds whenever every intersection of k of the sets A_j is empty ($k > 1$). In what follows, K and K' will always denote sets of k suffixes j ($1 \leqslant j \leqslant n$) for which the corresponding intersections A_K, $A_K{}'$, are not empty; J, J', etc. denote (as hitherto) arbitrary non-null sets of suffixes; and, except where the contrary is stated, all suffixes and superscripts run over all their admissible values.

From the definition of k and the properties of a d.s., we have

(3) $$A_K{}^\beta \cap A_J{}^\alpha = 0 \quad \text{unless } J \subset K \text{ and } A_J{}^\alpha \supset A_K{}^\beta.$$

In particular, the sets $A_K{}^\beta$ are all pairwise disjoint. Hence there exist open sets $V(K, \beta)$ such that

(4) $$\begin{aligned} &A_K{}^\beta \subset V(K, \beta),\\ &V(K, \beta) = 0 \text{ whenever } A_K{}^\beta = 0,\\ &\mathrm{Cl}(V(K, \beta)) \subset U(J, \alpha) \text{ whenever } A_K{}^\beta \subset A_J{}^\alpha,\\ &\mathrm{Cl}(V(K, \beta)) \cap A_J{}^\alpha = 0 \text{ whenever } A_K{}^\beta \cap A_J{}^\alpha = 0, \text{ and}\\ &\mathrm{Cl}(V(K, \beta)) \cap \mathrm{Cl}(V(K', \beta')) = 0 \text{ unless } K = K' \text{ and } \beta = \beta'. \end{aligned}$$

From (1) and (2), we may further suppose that each $V(K, \beta)$ is an open F_σ and is connected relative to $A_K{}^\beta$.

Now write

(5) $$W = \bigcup V(K, \beta),\ A'_j = A_j - W,\ A'_{J}{}^{\alpha} = A_J{}^{\alpha} - W.$$

Clearly the sets A'_j are closed, $\{A'_J{}^{\alpha}\}$ is a d.s. of $\{A'_j\}$, and no k of the sets A'_j have a common point. Again, in view of (3), there exist open sets $U'(J,\alpha)$ such that

(6) $$\begin{gathered} A'_J{}^{\alpha} \subset U'(J, \alpha) \subset U(J, \alpha), \\ \mathrm{Cl}(U'(J, \alpha)) \cap \mathrm{Cl}(V(K, \beta)) = 0 \text{ whenever } A_K{}^{\beta} \cap A_J{}^{\alpha} = 0, \text{ and} \\ \mathrm{Cl}(U'(J, \alpha)) \cap A_{J'}{}^{\alpha'} = 0 \text{ whenever } A_J{}^{\alpha} \cap A_{J'}{}^{\alpha'} = 0. \end{gathered}$$

Applying the hypothesis of induction to the system $\{A'_J\}$ and open sets $U'(J, \alpha)$, we obtain open F_σ sets $B'_1, \dots, B'_n$, with a d.s. $\{B'_J{}^{\alpha}\}$, having the properties corresponding to (i)—(v) of the lemma. Define

(7) $$\begin{aligned} B_j &= B'_j \cup \bigcup\{V(K, \beta) \mid j \in K\}, \text{ and} \\ B_J{}^{\alpha} &= B'_J{}^{\alpha} \cup \bigcup\{V(K, \beta) \mid A_K{}^{\beta} \cap A_J{}^{\alpha} \neq 0\} \\ &= B'_J{}^{\alpha} \cup \bigcup\{V(K, \beta) \mid K \supset J \text{ and } A_K{}^{\beta} \subset A_J{}^{\alpha}\} \end{aligned}$$

(as follows from (3) and (4)). Clearly B_j is an open F_σ, and $B_J{}^{\alpha}$ is connected relative to A_J (for $B'_J{}^{\alpha}$ is connected relative to $A'_J{}^{\alpha} \subset A_J{}^{\alpha}$, and each $V(K, \beta)$ occurring is connected relative to $A_K{}^{\beta} \subset A_J{}^{\alpha}$). It follows easily from (4), (6), and the hypothesis of induction that $\mathrm{Cl}(B_J{}^{\alpha}) \subset U(J, \alpha)$. To prove $A_J{}^{\alpha} \subset B_J{}^{\alpha}$, suppose $x \in A_J{}^{\alpha} - B_J{}^{\alpha}$; then x non $\in A'_J{}^{\alpha}$ (else $x \in B'_J{}^{\alpha} \subset B_J{}^{\alpha}$), and so, from (5), $x \in W$, say $x \in V(K, \beta)$. From (4), $A_K{}^{\beta} \cap A_J{}^{\alpha} \neq 0$; hence, from (7), $V(K, \beta) \subset B_J{}^{\alpha}$, contradicting x non $\in B_J$. Thus properties (i) and (ii) are established.

Property (iii) is proved as follows. Suppose $A_J{}^{\alpha} \cap A_{J'}{}^{\alpha'} = 0$ and

$$x \in \mathrm{Cl}(B'_J{}^{\alpha} \cup V(K, \beta)) \cap \mathrm{Cl}(B'_{J'}{}^{\alpha'} \cup V(K', \beta')),$$

where (from (7)) $K \supset J$, $K' \supset J'$, $A_K{}^{\beta} \subset A_J{}^{\alpha}$ and $A_{K'}{}^{\beta'} \subset A_{J'}{}^{\alpha'}$; we must derive a contradiction. The hypothesis of induction gives $\mathrm{Cl}(B'_J{}^{\alpha}) \cap \mathrm{Cl}(B'_{J'}{}^{\alpha'}) = 0$, while from (4) we obtain $\mathrm{Cl}(V(K, \beta)) \cap \mathrm{Cl}(V(K', \beta')) = 0$. Hence we may assume

$$x \in \mathrm{Cl}(V(K, \beta)) \cap \mathrm{Cl}(B'_{J'}{}^{\alpha'}) \subset \mathrm{Cl}(V(K, \beta)) \cap \mathrm{Cl}(U'(J', \alpha')).$$

From (6) we must have $A_K{}^{\beta} \cap A_{J'}{}^{\alpha'} \neq 0$, and therefore (from (3)) $A_K{}^{\beta} \subset A_{J'}{}^{\alpha'}$. But this contradicts the assumption $A_J{}^{\alpha} \cap A_{J'}{}^{\alpha'} = 0$.

Property (iv) is immediate from (7), (5) and the hypothesis of induction. Thus all that remains to be proved, apart from (v), is that $\{B_J{}^{\alpha}\}$ is in fact a d.s. of $\{B_j\}$; and in virtue of (iii) and (iv) it will suffice to verify that

(8) $$\bigcup_\alpha B_J{}^{\alpha} = B_J, \text{ where } B_J = \bigcap\{B_j \mid j \in J\}.$$

First suppose $x \in B_J{}^{\alpha}$. If $x \in B'_J{}^{\alpha}$, then $x \in \bigcap\{B'_j \mid j \in J\} \subset B_j$; hence we may suppose $x \in V(K, \beta)$ where (from (7)) $A_K{}^{\beta} \subset A_J{}^{\alpha}$ and $K \supset J$. Thus (7) gives $V(K, \beta) \subset B_j$ whenever $j \in J$, so again $x \in B_J$. This proves $\bigcup_\alpha B_J{}^{\alpha} \subset B_J$.

Conversely, suppose $x \in B_J$. If for every $j \in J$ we have $x \in B'_j$, then

$x \in B'_J = \bigcup_\alpha B'_J{}^\alpha \subset \bigcup_\alpha B_J{}^\alpha$, as desired. Thus we may assume (from (7)) that $x \in V(K, \beta)$, where $j \in K$, for at least one $j \in J$. We assert $J \subset K$. For if say $j' \in J - K$, then $x \in B'_{j'}$, since otherwise $x \in V(K', \beta')$ with $j' \in K'$, and then $V(K', \beta') \cap V(K, \beta) \neq 0$ though $K \neq K'$, contradicting (4). Thus $x \in \bigcup_\gamma B'_{j'}{}^\gamma \subset \bigcup U'(j', \gamma)$, and so for some γ we have $x \in U'(j', \gamma) \cap V(K, \beta)$, which from (6) implies $A_K{}^\beta \cap A_{j'}{}^\gamma \neq 0$, whence (by (3)) $j' \in K$, a contradiction. Thus $J \subset K$; and the definition of a d.s. now gives the existence of an α' such that $A_K{}^\beta \subset A_J{}^{\alpha'}$. From (7), we have $V(K, \beta) \subset B_J{}^{\alpha'}$, and so $x \in \bigcup_\alpha B_J{}^\alpha$, completing the proof of (8).

Finally, the verification of (v) is along similar lines, and is left to the reader.

6.2. *Strictly canonical mappings.* Let $U_1, U_2, \ldots, U_n$ be a given covering of S, with a given d.s. $\mathfrak{D} = \{U_J{}^\alpha\}$. For each $x \in S$, let $J(x)$ be the set of all suffixes j for which $x \in U_j$; thus $x \in U_{J(x)}$, and so $x \in U_{J(x)}{}^\alpha$ for one and only one value of α, say for $\alpha = \alpha(x)$. The corresponding (open) simplex $u_{J(x)}{}^{\alpha(x)}$ of $\mathfrak{M}(\mathfrak{D})$ will be denoted by $\sigma(x)$.

A continuous mapping h of S in $\mathfrak{M}(\mathfrak{D})$ will be called *strictly canonical*[11] if it satisfies

(1) $$h(x) \in \sigma(x), \quad \text{all } x \in S.$$

It is easy to see that (1) is equivalent[12] to

(2) $$h^{-1}(\mathrm{St}\, u_J{}^\alpha) = U_J{}^\alpha,$$

St $u_J{}^\alpha$ denoting the (open) star of the simplex $u_J{}^\alpha$ in $\mathfrak{M}(\mathfrak{D})$.

The proof of the standard existence theorem for mappings in ordinary nerves can readily be extended to give:

LEMMA 2. *Let $U_1, U_2, \ldots, U_n$ be open F_σ sets which cover S and let $\mathfrak{D} = \{U_J{}^\alpha\}$ be a d.s. of $\{U_j\}$. Then there exists a strictly canonical mapping h of S in $\mathfrak{M}(\mathfrak{D})$.*

6.3. The fundamental lemma is the following analogue of a lemma of Eilenberg [4, p. 105], and the idea of the proof is essentially the same, though with some complications.

LEMMA 3. *Let $B_1, B_2, \ldots, B_n$ be a covering of S by open F_σ sets of finite incidence, with $\{B_J{}^\alpha\}$ as natural d.s., and suppose that $\{\mathrm{Cl}(B_J{}^\alpha\}$ is a d.s. of $\{\overline{B_j}\}$. Let h be a strictly canonical mapping of S in the modified nerve $\mathfrak{M}$ of $\{B_j\}$, and let f be a mapping of $\mathfrak{M}$ in S^1 such that $fh \sim 1$ on S. Then $f \sim 1$ on $\mathfrak{M}$.*

Suppose not. Then, as in [4, p. 105], there exists a simple closed edge-path in $\mathfrak{M}$ on which f non ~ 1; let $\mathfrak{C}$ be such a closed edge-path having as few edges as possible. There is no loss of generality in assuming the sets B_j to be connected (otherwise we replace them by their components); hence the nota-

[11] Compare [1, p. 210].

[12] For ordinary nerves it is enough to require that (2) hold for vertex-stars; but this reduction is no longer valid for modified nerves, in general.

tion may be chosen so that $\mathfrak{C}$ consists of the edges $b_{12}^1, b_{23}^1, \ldots, b_{(s-1)s}^1, b_{s1}^1$ joining successive vertices $b_1, b_2, \ldots, b_s$. (Note that here s may well equal 2.) As in [4], it follows from 2.2(1) and the choice of $\mathfrak{C}$ that $B_j \cap B_k = 0$ $(1 \leqslant j < k \leqslant s)$ unless j, k are consecutive in the cyclic order $12 \ldots$ sl; and thence it follows, if $s > 3$, that no three of the sets B_j $(1 \leqslant j \leqslant s)$ can have a common point. Further, this holds even if $s = 3$. For otherwise b_1, b_2, b_3 are the vertices of a 2-cell b_{123}^α in $\mathfrak{M}$, which will have edges say $b_{23}^\beta, b_{31}^\gamma, b_{12}^\delta$; but $f \sim 1$ on $\mathrm{Cl}(b_{123}^\alpha)$ (from 2.2(4)), and also $f \sim 1$ on $\mathrm{Cl}(b_{23}^1) \cup b_{23}^\beta$ (which is either an arc or a closed edge-path shorter than $\mathfrak{C}$), and similarly $f \sim 1$ on $\mathrm{Cl}(b_{31}^1) \cup b_{31}^\gamma$ and on $\mathrm{Cl}(b_{12}^1) \cup b_{12}^\delta$, so that (from 2.2(1)) $f \sim 1$ on $\mathfrak{C}$, which is absurd. Hence, in view of the postulates on the sets B_j, we have:

(1) No three of the sets $\overline{B}_j$ have a common point $(1 \leqslant j \leqslant s)$.

Write $S' = \bigcup B_j$ $(1 \leqslant j \leqslant s)$; evidently S' is connected, and further, as an open F_σ subset of S, S' is also locally connected and normal. In the next paragraph, all considerations will be relative to S', and we use dashes to indicate relative closures and frontiers. The suffixes j, k, will run between 1 and s, and will be taken modulo s.

For each fixed j we have

$$\mathrm{Fr}'(B_j) \subset \mathrm{Cl}'(B_j) \cap \mathrm{Cl}'(B_{j-1} \cup B_{j+1}) = \textstyle\bigcup_\alpha \mathrm{Cl}'(B_{(j-1)j}^\alpha) \cup \bigcup_\beta \mathrm{Cl}'(B_{j(j+1)}^\beta),$$

the union of a finite number of pairwise disjoint and (relatively) closed connected non-empty sets. On applying [10, 3.4] in S', we obtain connected sets $H_j^\alpha \supset \mathrm{Cl}'(B_{(j-1)j}^\alpha)$, $K_j^\beta \supset \mathrm{Cl}'(B_{j(j+1)}^\beta)$, no three of which have a common point (j being fixed), such that the intersection of every two of these sets is contained in $B_j - \bigcup\{\mathrm{Cl}'(B_k) | k \neq j\}$. Moreover, the sets H_j^α, K_j^β, so obtained will in the first instance satisfy $\bigcup_\alpha H_j^\alpha \cup \bigcup_\beta K_j^\beta = \mathrm{Cl}'(B_j)$, and will be closed (relative to S'); but we replace them (using 6.1(1) and 6.1(2)) by slightly larger sets to make them open F_σ's (relative to S' and thus also relative to S) without introducing any further intersections. For convenience, we introduce the symbol L_j^α to stand for either H_j^α or K_j^α. If now j is allowed to vary, we see that, while $H_j^\alpha \cap K_{j-1}^\alpha \supset B_{(j-1)j}^\alpha \neq 0$, all other intersections of the form $L_j^\alpha \cap L_k^\beta$ $(j \neq k)$ are empty, and consequently no three of the sets L_j^α, $1 \leqslant j \leqslant s$, can have a common point.

Let $\mathfrak{N}$ denote the (unmodified) nerve of the sets L_j^α; clearly $\mathfrak{N}$ is a linear graph. We use $h_j^\alpha, k_j^\alpha, l_j^\alpha$ for the vertices of $\mathfrak{N}$ corresponding to the sets H_j^α, K_j^α, L_j^α respectively. Since B_j is connected, there exists a simple edge-path C_j in $\mathfrak{N}$, joining h_j^1 to k_j^1 via vertices of the form l_j^α (j fixed) only; and since $B_{j(j-1)}^1 \neq 0$ there exists a 1-cell $(k_{j-1}^1 h_j^1)$ in $\mathfrak{N}$. The sequence

$$\mathfrak{K} = (k_s^1 h_1^1),\ C_1,\ (k_1^1 h_2^1),\ C_2, \ldots, (k_{s-1}^1 h_s^1),\ C_s,$$

constitutes a simple closed curve in $\mathfrak{N}$.

Now consider the (continuous) simplicial mapping ϖ of $\mathfrak{N}$ in $\mathfrak{M}$ defined as follows: ϖ maps each vertex l_j^α and edge $l_j^\alpha l_j^\beta$ on the vertex b_j of $\mathfrak{M}$, and maps

each edge of the form $k_{j-1}{}^*h_j{}^*$ "linearly" on the edge $b_{(j-1)j}{}^*$ of $\mathfrak{M}$. Clearly $\varpi(C_j) = b_j$ and so ϖ maps $\mathfrak{K}$ on $\mathfrak{C}$ with degree 1. From this and the uniform continuity of f, we obtain a sequence of mappings $\varpi = \varpi_0, \varpi_1, \ldots, \varpi_\mu$ of $\mathfrak{K}$ on $\mathfrak{C}$ such that

(i) ϖ_μ is a homeomorphism of $\mathfrak{K}$ on $\mathfrak{C}$,

(ii) $|f(\varpi_{\lambda-1}(x)) - f(\varpi_\lambda(x))| < 1$ for all $x \in \mathfrak{K}$ $\quad (1 \leqslant \lambda \leqslant \mu)$.

Thus, from 2.2(3), $f\varpi \sim f\varpi_1 \sim \ldots f\varpi_\mu$ on $\mathfrak{K}$; and from the fact that f non ~ 1 on $\mathfrak{C}$, we readily deduce $f\varpi_\mu$ non ~ 1 on $\mathfrak{K}$, and consequently $f\varpi$ non ~ 1 on $\mathfrak{K}$.

Thus there exist simple closed edge-paths in $\mathfrak{N}$ on which $f\varpi$ non ~ 1; let $\mathfrak{K}_0$ be one having as few edges as possible, and let the corresponding sets $L_j{}^*$ be renamed $L(1), L(2), \ldots, L(p), L(1)$, following the cyclic order of $\mathfrak{K}_0$. (Note that now $p \geqslant 3$.) As before, two sets $L(j), L(k)$ meet if and only they are consecutive in this cyclic order; hence the nerve of $L(1), L(2), \ldots, L(p)$ is precisely $\mathfrak{K}_0$. Write $Q = \bigcup L(j)$ $(1 \leqslant j \leqslant p)$, and let h' be a strictly canonical mapping of Q in $\mathfrak{K}_0$. It is easy to see that, for each $x \in Q$, the point $\varpi h'(x)$ of $\mathfrak{M}$ belongs to the closure of the simplex $\sigma(x)$ of $\mathfrak{M}$ which contains $h(x)$. Let $h(x) = h_0(x),\ h_1(x), \ldots, h_N(x) = \varpi h'(x)$ be points dividing the "straight" segment joining $h(x)$ to $\varpi h'(x)$, in $\mathrm{Cl}(\sigma(x))$, into N equal parts. One readily verifies that each h_k is a continuous mapping of Q in $\mathfrak{M}$, and that, from 2.2(3), $fh_0 \sim fh_1 \sim \ldots \sim fh_N$ if N is large enough. Thus $f\varpi h' \sim fh \sim 1$ on Q.

The argument can be concluded as in [4, p. 106]; alternatively, by the theorem there proved, we must have $f\varpi \sim 1$ on $\mathfrak{K}_0$. contradicting the definition of $\mathfrak{K}_0$.

6.4. THEOREM 6. *Let $A_1, A_2, \ldots, A_n$ be non-empty closed connected sets of finite incidence which cover S; let $\mathfrak{N}$ be their nerve and $\mathfrak{M}$ their modified nerve. Then*

$$r(\mathfrak{N}) \leqslant r(\mathfrak{M}) \leqslant r(S).$$

That $r(\mathfrak{N}) \leqslant r(\mathfrak{M})$ has been proved in 1.3 Suppose $r(\mathfrak{M}) \geqslant m$; from Theorem 5 (5.1) it will suffice to prove $\rho(S) \geqslant m$. There exist closed subsets M, N of $\mathfrak{M}$, and m independent mappings f_j $(1 \leqslant j \leqslant m)$ of $\mathfrak{M}$ in S^1, such that $M \cup N = \mathfrak{M}$, $f_j \sim 1$ on M, and $f_j \sim 1$ on N. By Lemma 1 (6.1), we can enlarge the sets A_j to open F_σ sets B_j having the same modified nerve $\mathfrak{M}$ and satisfying the hypotheses of Lemma 3 (6.3). By Lemma 2(6.2), there exists a strictly canonical mapping h of S in $\mathfrak{M}$. let $X = h^{-1}(M)$ and $Y = h^{-1}(N)$; X and Y are closed sets covering S, and each of the m mappings f_jh of S in S^1 evidently satisfies $f_jh \sim 1$ on X and $f_jh \sim 1$ on Y. But Lemma 3 (6.3) shows that these mappings are independent on S; hence $\rho(S) \geqslant m$, and the theorem is proved.

6.5. THEOREM 7. *Let $A_1, A_2, \ldots, A_n$ be non-empty, connected, locally connected, normal sets of finite incidence, which cover S and are such that $A_j - A_k$ and $A_k - A_j$ are always separated*[6]. *Let $\mathfrak{M}$ be their modified nerve. Then $r(S) \leqslant b_1(\mathfrak{M}) + \sum r(A_j)$.*

We may assume $r(A_j) = r_j < \infty$. Suppose there exist N independent mappings $f_1, f_2, \ldots, f_N$ of S in S^1, and closed sets X, Y such that $X \cup Y = S$, $f_j \sim 1$ on X, and $f_j \sim 1$ on $Y (1 \leqslant j \leqslant n)$; we must prove (in view of Theorem 5, 5.1) that $N \leqslant b_1(\mathfrak{M}) + \sum r_j$.

Since $f_j \sim 1$ on $X \cap A_1$ and on $Y \cap A_1$, Theorem 5 shows that at most r_1 of the mappings f_j can be independent on A_1. Let the greatest number of independent mappings f_j on A_1 be $s_1 \leqslant r_1$; we may suppose the notation so chosen that $f_1, \ldots, f_{s_1}$ are independent on A_1, and obtain for each $j > s_1$ a relation, say

$$g_j \equiv f_j^{p_j} f_1^{q_{j1}} \ldots f_{s_1}^{q_{js_1}} \sim 1 \text{ on } A_1,$$

where the exponents are integers and clearly $p_j \neq 0$. It readily follows that the $N - s_1$ mappings g_j are independent on S, and satisfy $g_j \sim 1$ on X and on Y.

By repeating this argument, applying it to $A_2, \ldots, A_n$ in turn, we obtain $N - \sum s_k$ independent mappings (say) h_j of S in S^1 (expressible as power-products of the N given mappings f_j), where $s_k \leqslant r_k$, such that $h_j \sim 1$ on each A_k $(1 \leqslant k \leqslant n)$. Hence, from 2.3(1),

$$N - \sum s_k \leqslant p(A_1, A_2, \ldots, A_n) \leqslant b_1(\mathfrak{M}),$$

and the theorem follows.

COROLLARY. *If further the sets A_j are closed and unicoherent, and no three of them have a common point, then $r(S) = r(\mathfrak{M})$.*

For Theorem 7 gives $r(S) \leqslant b_1(\mathfrak{M}) = r(\mathfrak{M})$, since $\mathfrak{M}$ is now a graph; and on the other hand Theorem 6 (6.4) gives $r(S) \geqslant r(\mathfrak{M})$.

6.6. It is natural to ask whether, in Theorem 7 above, the term $b_1(\mathfrak{M})$ can be replaced by $r(\mathfrak{M})$. The answer is negative, as is shown by the following example: Let T be a 2-manifold of genus k, simplicially subdivided, and let $B_1, B_2, \ldots, B_n$ denote the closed stars of the vertices of T in the barycentric subdivision. Let C be a small circular region interior to B_1, and define $S = T - C$, $A_1 = B_1 - C$, and $A_j = B_j$ $(j \geqslant 2)$. It follows immediately from known theorems that $r(S) = 2k$, $r(A_1) = 1$, and $r(A_j) = 0$ $(j \geqslant 2)$. But the modified nerve $\mathfrak{M}$ of $A_1, A_2, \ldots, A_n$ is simply the nerve of $B_1, B_2, \ldots, B_n$—i.e., is T. Hence $r(\mathfrak{M}) = r(T) = k$.

However, the replacement of $b_1(\mathfrak{M})$ by $r(\mathfrak{M})$ in Theorem 7 is justified (under reasonable conditions) provided all the sets A_j are unicoherent. For simplicity we consider only the polyhedral case (though the generalization to ANR's would be easy), and in stating the result do not distinguish between "complex" and "polytope".

THEOREM 8. *Let $A_1, A_2, \ldots, A_n$ be closed, connected, non-empty unicoherent subcomplexes of a complex S, which cover S, and let $\mathfrak{M}$ be their modified nerve. Then $r(S) = r(\mathfrak{M})$.*

Sketch of proof. Choose points $p_J{}^\alpha \in A_J{}^\alpha$, $\{A_J{}^\alpha\}$ being the natural d.s. of $\{A_j\}$, and for each pair $A_J{}^\alpha$, $A_K{}^\beta$ with $K \supset J$ and $A_J{}^\alpha \supset A_K{}^\beta$, join $p_K{}^\alpha$ to $p_J{}^\beta$ by an arc in $A_J{}^\alpha$. These arcs form a graph G. There is an obvious mapping ϕ of the edge-paths in $\mathfrak{M}$ onto paths in $G \subset S$. In general, ϕ need not induce a homomorphism of $\pi_1(\mathfrak{M})$. However, if $r(S) = r$, there exists [4, p. 110] a homomorphism ψ of $\pi_1(S)$ onto F_r, the free (non-abelian) group on r generators. Using the fact that the sets A_j are unicoherent, one can show that $\psi\phi$ induces a homomorphism of $\pi_1(\mathfrak{M})$ onto F_r. Hence [4, p. 110] $r(\mathfrak{M}) \geqslant r$. But $r(\mathfrak{M}) \leqslant r$, by Theorem 6 (6.4); and Theorem 8 is established.

BIBLIOGRAPHY

[1] C. H. Dowker, *Mapping theorems for non-compact spaces*, Amer. J. Math., vol. 69 (1947), pp. 200-242.

[2] S. Eilenberg, *Transformations continues en circonférence et la topologie du plan*, Fund. Math., vol. 26 (1936), pp. 61-112.

[3] ———, *Sur les espaces multicohérents I*, Fund. Math., vol. 27 (1936), pp. 153-190.

[4] ———, *Sur les espaces multicohérents II*, Fund. Math., vol. 29 (1937), pp. 101-122.

[5] ———, *Sur la multicohérence des surfaces closes*, Comptes Rendus de l' Académie des Sciences de Warsaw, vol. 30 (1937), pp. 109-111.

[6] S. Lefschetz, *Algebraic Topology*, Amer. Math. Soc. Colloquium Publications, vol. 27 (1942).

[7] ———, *Topics in Topology*, Ann. of Math. Studies No. 10, Princeton, 1942.

[8] H. Seifert and W. Threlfall, *Lehrbuch der Topologie* (Leipzig, 1934).

[9] A. H. Stone, *Incidence relations in unicoherent spaces*, Trans. Amer. Math. Soc., vol. 65 (1949), pp. 427-447.

[10] ———, *Incidence relations in multicoherent spaces, I*, Trans. Amer. Math. Soc., vol. 66 (1949), pp. 389-406.

[11] A. D. Wallace, *Separation spaces*, Ann. of Math., vol. 42 (1941), pp. 687-697.

[12] G. T. Whyburn, *Analytic Topology*, Amer. Math. Soc. Colloquium Publications, vol. 28 (1942).

Manchester University

SOME PROPERTIES OF C-CONVEX SETS

F. A. VALENTINE

1. Introduction. The notion of convexity in $\mathfrak{R}_m$ (m-dimensional Euclidean space) can be generalized to apply to non-connected sets as follows.

DEFINITION 1. *A set is said to be C-convex if each of its components is convex. If the number of components of such a set is n, it is called a C_n-convex set.*

In order to determine the character of the complement of a C_n-convex set, we use the notion of L_n set, a concept studied by my colleague Alfred Horn and myself [2]. Although my original goal was to establish the fact that in the plane the complement of a bounded open C_n-convex set ($n > 1$) is an L_{n+1} set, the auxiliary concept of "Maximal families of disjoint open convex sets" almost preempted my original intention. For this reason, the latter concept has been studied in §3 separately. In order to complete the terminology, I restate the definition given by Horn and myself [2].

DEFINITION 2. *A set S is called an L_n set if each pair of points in S can be joined by a polygonal arc in S having at most n segments.*

Throughout this paper we confine ourselves to sets in $\mathfrak{R}_2$.

2. Polygonal sets in the plane. In the following treatment the words vertex, edge and face are used in the usual sense [3, pp. 194-5]. An edge is always incident with a face, and a face may be bounded or unbounded. A linear edge is one which is contained in a straight line.

DEFINITION 3. *A polygonal set P_n is a connected closed set which has the following properties.*

(a) *It is the sum of a finite number of linear edges.*

(b) *Its complement consists of n components, and each of these is convex (called a face).*

(c) *Each vertex of P_n is incident with at least three edges.*

NOTATION. A polygonal arc P in P_n joining x and y is denoted by $xx_1 \ldots x_ty$, where $x_1, \ldots, x_t$ denote the vertices of P_n on P distinct from x and y. If no such vertices exist, then $P = xy$. The boundary of a face F of P_n is denoted by $B(F)$.

Received September 4, 1949. Presented to the American Mathematical Society, April 30, 1949.

DEFINITION 4. *An improper vertex of P_n is one which is incident with at least four edges of P_n. A segment of a polygonal arc P in P_n, as distinguished from an edge of P_n, is a maximal connected linear subset of P.*

DEFINITION 5. *If a polygonal arc in P_n joining x and y has a shortest length (a proper or improper minimum) relative to the arcs in P_n joining x and y, it is called a minimal polygonal arc, and we denote it by $P(x, y)$.*

LEMMA 1. *Let F be a face of P_n. If $x \in B(F)$, $y \in B(F)$, then any minimal polygonal arc $P(x, y) \subset B(F)$.*

Lemma 1 is an immediate consequence of the convexity of F.

LEMMA 2. *Let $P(x, y) = xx_1 \ldots x_t y$ be a minimal polygonal arc in P_n. Let $\mathfrak{F}_i = (F_{i1}, F_{i2}, \ldots, F_{im_i})$ denote the collection of faces of P_n which have x_i as a vertex, and which do not have $x_{i-1}x_i$ as an edge ($i = 1, \ldots, t$; $x_0 = x$). Then all of the faces in the collection $\sum_{i=1}^{t} \mathfrak{F}_i$ are distinct.*

Proof. Condition (c) in Definition 3 implies that $m_i \geqslant 1$ ($i = 1, \ldots, t$). Suppose there exist two faces F_{is} and F_{kr} contained in $\sum_{i=1}^{t} \mathfrak{F}_i$ such that $F_{is} = F_{kr}$ ($1 \leqslant i < k \leqslant t$). By Lemma 1, we have then $P(x_i, x_k) = x_i x_{i+1} \ldots x_k \subset B(F_{kr})$. However, since by hypothesis, $x_{k-1}x_k \not\subset B(F_{kr})$, we have $x_{k-1}x_k \not\subset P(x, y)$, which is a contradiction. Hence Lemma 2 is clearly true.

THEOREM 1. *Let $P(x, y) = xx_1 \ldots x_t y$ be a minimal polygonal arc in P_n. Then there exists a collection $\mathfrak{F} = (F_0, F_1, F_2, \ldots, F_t)$ of distinct faces of P_n such that the edge $x_i x_{i+1} \subset B(F_i)$ ($i = 0, \ldots, t$; $x_0 = x$, $x_{t+1} = y$). Let p denote the number of faces in $\mathfrak{G} = \sum_{i=1}^{t} \mathfrak{F}_i - \mathfrak{F}$, and let v be the number of faces in P_n not incident with any part of $P(x, y)$. Then $p + t + v \leqslant n - 2$.*

Proof. Theorem 1 follows from Lemma 2. Let F_0 and F'_0 be the faces of P_n incident with xx_1. As in the proof of Lemma 2, F_0 non $\in \sum_{i=1}^{t} \mathfrak{F}_i$, F'_0 non $\in \sum_{i=1}^{t} \mathfrak{F}_i$, since $P(x, y)$ is minimal. Define F_k to be a member of $\mathfrak{F}_k$ having $x_k x_{k+1}$ as an edge ($k = 1, \ldots, t$). Hence $\mathfrak{F}$ has been defined, and it contains distinct members. Moreover, since F'_0 non $\in \mathfrak{F}$, F'_0 non $\in \mathfrak{G}$, by counting distinct faces, we get $p + t + 1 + v \leqslant n - 1$.

COROLLARY 1. *A polygonal set P_n ($n \geqslant 2$) is an L_{n-1} set.*

3. Maximal families of convex sets in the plane.

DEFINITION 6. *A family of disjoint open convex sets is said to be maximal if no member of the family is a proper subset of an open convex set which is disjoint with the rest of the family.*

A family of this type containing exactly n members is called an M_n set.

LEMMA 3. *Each member of an open C_n-convex set ($n > 1$) can be enclosed in an open convex set which has a polygonal boundary, and which is disjoint with the rest of C_n. The boundary of this set need not be connected.*

This lemma was proved by Stoelinga. See Bonnesen and Fenchel [1, p. 5].

THEOREM 2. *The boundary of a maximal family M_n ($n > 1$) of disjoint open convex sets is the sum of a finite number of line segments, lines and half-lines.*

Proof. Each member of M_n must be a two-dimensional convex plane polygon, otherwise by Lemma 3, it would not be maximal. Since there are a finite number of members in M_n, each of which has a finite number of linear elements in its boundary, the boundary of M_n is the sum of a finite number of line segments, lines and half-lines.

DEFINITION 7. *A component of the complement (face) of a polygonal set is called a pinwheel R provided:*

(i) *It is a bounded convex set.*

(ii) *The vertices of $\overline{R}$ can be ordered consecutively ($x_1, x_2, \ldots, x_t; x_t = x_1$) so that for each vertex x_i there exists an edge E_i of the polygonal set which abuts R externally at x_i, and which is a linear extension of $x_{i-1}x_i$ ($i = 2, \ldots, t$).* (See Figure 1; $E_t = E_1$).

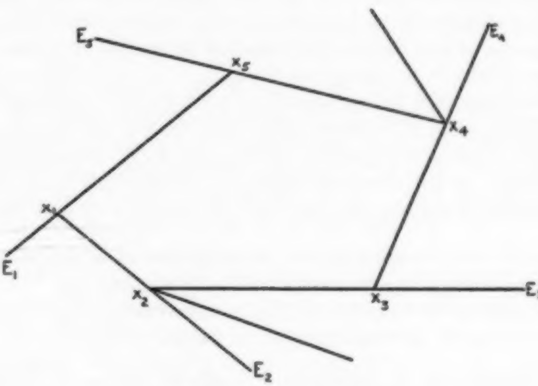

FIGURE 1. A pinwheel

THEOREM 3. *Each component of the complement of the closure of a maximal family M_n is a pinwheel.*

Proof. Let K be any component of the complement of $\overline{M}_n$. By Theorem 2, K has a boundary consisting solely of line segments, lines or half-lines. Let $B(K)$ be a *component* of the boundary of K. Among the finite number of vertices of $B(K)$ we include corners as well as vertices of the boundary of M_n. Since M_n is maximal, there exists a finite edge $x_1x_2 \subset B(K)$.

Since M_n contains a finite number of components, let C_j be the member of M_n abutting x_1x_2. The straight line through x_1x_2 determines two open half-

planes $\Re_1^+$ and $\Re_1^-$ where $C_j \subset \Re_1^+$ by definition. Let E_1 and E_2 be the edges of C_j which abut K at x_1 and x_2 respectively. Since M_n is a maximal family of convex sets, $E_1+E_2-x_1-x_2 \not\subset \Re_1^+$. Moreover, since C_j is convex, $E_1+E_2-x_1-x_2 \not\subset \Re_1^-$. Hence at least one of the edges E_1 and E_2 is a linear extension of x_1x_2. Without loss of generality suppose E_2 is an extension of x_1x_2. Hence, x_2 must be a vertex of the boundary of M_n, so that at least *three* edges of the boundary of M_n are incident with x_2. Hence, the interior angle θ_2 of $\overline{K}$ at x_2 is less than π. Let x_2x_3 denote the edge (finite or infinite) of $\overline{K}$ which together with x_1x_2 makes the angle θ_2. The edge x_2x_3 must be *finite*, otherwise the member of M_n abutting x_2x_3 would not be maximal relative to M_n. By induction, we get a finite polygonal line $x_1x_2 \ldots x_s$ and a set of extensions $E_i (i = 2, \ldots, s)$ such that the interior angle of $\overline{K}$ at x_i is less than π, and such that E_i is an extension of $x_{i-1}x_i$. Since $B(K)$ has a finite number of vertices including corners, it is clear that this sequence $x_1x_2 \ldots x_s$ can only be continued until we get $x_1x_2 \ldots x_{t-1}x_t$ where $x_1, \ldots x_{t-1}$ are all distinct, and where x_t is one of the vertices $x_1, x_2, \ldots, x_{t-2}$. One can prove that $x_t = x_1$, otherwise all the extensions E_i would not exist, which is a contradiction. Since M_n is maximal, the interior angle of the simple closed polygon $x_1x_2 \ldots x_t$ $(x_t = x_1)$ at x_1 is also less than π. Hence $x_1x_2 \ldots x_t$ is a closed convex polygonal curve. Since K is connected, $B(K)$ is contained in the closed convex set bounded by $x_1x_2 \ldots x_t$, and it follows by an argument of the type just given for $x_1x_2 \ldots x_t$ that $B(K) = x_1x_2 \ldots x_t$.

Finally, we show that the set bounded by $B(K)$ is K. Suppose a component $B_1(K)$ of the boundary of K exists which is interior to the convex set bounded by $B(K)$. By virtue of the previous paragraph, $B_1(K)$ would bound a convex set, at least part of which would belong to K. But this would make K disconnected, which is a contradiction. Thus K satisfies (i) and (ii).

THEOREM 4. *Each component of the boundary of a maximal family M_n $(n > 1)$ of disjoint open convex convex sets is a polygonal set.*

The complement of the boundary of M_n is a maximal family M_r $(r \geqslant n)$, where $r - n$ is the number of pinwheels in the complement of $\overline{M}_n$.

Proof. Property (a) in Definition 3 holds by virtue of Theorem 2. Properties (b) and (c) hold since each member of M_n is convex, and since each residual domain of $\overline{M}_n$ is convex. The concluding statement follows from Theorem 3.

THEOREM 5. *The boundary of a maximal family M_n $(n > 1)$ has α components if and only if $\alpha - 1$ members of M_n are slabs (A slab is an open convex set bounded by two parallel lines).*

Proof. Let T be a component of the boundary of M_n. The set T must be unbounded, otherwise the unbounded component of M_n abutting T externally would not be convex. If the boundary of each member of M_n incident[1] with T

[1]A member of M_n is said to be incident with T if its boundary contains at least one edge of T.

is connected, then the boundary of M_n is in T, and T is the only component of the boundary of M_n. *If a member of M_n has a disconnected boundary, then it must be a slab, since it is convex.* The set T can have at most two slabs abutting it, since two disjoint slabs must be parallel. All the slabs in M_n then must be parallel, and between two consecutive slabs there can be at most one component of the boundary of M_n. These facts clearly imply the conclusions of Theorem 5.

In the following treatment it should be recalled that in the definition of an L_n set, the word segment was used, and not edge (See Definitions 2 and 4).

THEOREM 6. *Let T be a component of the boundary of a maximal family M_n ($n > 1$) of disjoint open convex sets. Let s be the number of members of M_n which are incident with T. Then T is an L_{s-1} set.*

Proof. Replace each slab abutting T (if any exist) by the half-plane which contains that slab, and which abuts T. The thus modified s sets of M_n incident with T form a maximal family M_s. The complement of T is a maximal family M_r. By Theorem 4, $r - s = q$ is the number of pinwheels in $M_r - M_s$. We designate the closures of these by $R_k (k = 1, \ldots, q)$. Choose $x \in T$, $y \in T$. If $P(x, y) = xy$, then it contains at most $s - 1$ segments. Let $P(x, y) = xx_1 \ldots x_t y$ and $\mathfrak{F}$ and $\mathfrak{G}$ denote the quantities described in Theorem 1.

Case 1. Suppose x non $\in R_k$, y non $\in R_k$ $(k = 1, \ldots, q)$. First, let S_β $(\beta = 1, \ldots, q_1)$ denote the closures of the pinwheels in $M_r - M_s$ each of which has one and only one vertex in common with $P(x, y) - x - y$. Since each of these vertices is then improper, we have $S_\beta \in \mathfrak{G}$ $(\beta = 1, \ldots, q_1)$. Set up an order on $P(x, y)$ from x to y, and let $Q_j (j = 1, \ldots, q_2)$ denote in succession the closures of the pinwheels in $M_r - M_s$ for which $Q_1 \cdot P(x, y)$ contains at least one edge of T. Each set $Q_j \cdot P(x, y)$ is connected, and $Q_1 \cdot P(x, y)$ precedes $Q_2 \cdot P(x, y)$, on $P(x, y)$ etc. Let x_j^1 and x_j^2 denote the vertices of T where $P(x, y)$ enters and leaves Q_j respectively. *If a vertex of T is an interior point of a segment of $P(x, y)$, it is called a removable vertex of $P(x, y)$.* If x_1^1 and x_1^2 are both proper vertices of Q_1, then since Q_1 is a pinwheel, either x_1^1 or x_1^2 is a removable vertex of $P(x, y)$. If either x_1^1 or x_1^2 is an improper vertex of Q_1, the set $\mathfrak{G}$ in Theorem 1 contains at least one face corresponding to that vertex. Hence Q_1 corresponds either to a face of $\mathfrak{G}$ or to a removable vertex of $P(x, y)$. If $x_1^2 \neq x_2^1$, then Q_1 is isolated from Q_2. If $x_1^2 = x_2^1$, then x_2^1 is improper. Moreover Q_1 and Q_2 then have opposite orientations in the sense that the vertices of one of them are ordered clockwise and the vertices of the other counterclockwise. (See Figure 1.) One can show that this implies the following. If x_2^1 is *not* a removable vertex of $P(x, y)$, then either x_1^1 or x_2^2 must be an improper vertex or a removable vertex of $P(x, y)$. This is true whether the sense in which the directed $P(x, y)$ meets Q_1 and Q_2 coincides with their proper orientations or not. Hence, we can assign to each Q_1 and Q_2 either a member of $\mathfrak{G}$ or a removable vertex of $P(x, y)$, and the faces and vertices involved are all distinct. Suppose Q_f, $Q_{f+1}, \ldots, Q_{f+\sigma}$

are a subset of consecutive sets from $Q_j (j = 1, \ldots, q_2)$ for which $x_f^2 = x_{f+1}^1$, $x_{f+1}^2 = x_{f+2}^1, \ldots, x_{f+\sigma-1}^2 = x_{f+\sigma}^1$. Then all of these vertices are improper. If none of these vertices is also a removable vertex of $P(x, y)$, then since each consecutive pair of $Q_f, Q_{f+1}, \ldots, Q_{f+\sigma}$ have opposite orientations, one can show that either x_f^1 or $x_{f+\sigma}^2$ is a removable vertex of $P(x, y)$ or an improper vertex. Hence, to each set in the above consecutive sets we can assign either a distinct face in $\mathfrak{G}$ or a removable vertex of $P(x, y)$. Moreover, one can choose the faces of $\mathfrak{G}$ just mentioned distinct from $\sum_{\beta=1}^{q_1} S_\beta$. *Now*, by separating $P(x, y) \cdot \sum_{j=1}^{q_2} Q_j$ into disjoint parts, the above type of argument implies the following. There is a subset of faces in $\mathfrak{G} - \sum_{\beta=1}^{q_1} S_\beta$ and a set of distinct removable vertices of $P(x, y)$ which together are in $1 - 1$ correspondence with $Q_1, Q_2, \ldots, Q_{q_2}$. Hence, if we let m equal the number of *segments* in $P(x, y)$, the above together with the fact $S_\beta \subset \mathfrak{G}$ $(\beta = 1, \ldots, q_1)$ implies that $m + q_1 + q_2 \leqslant t + 1 + p$. Theorem 1 implies that $p + t + 1 + v \leqslant r - 1$. Since $q_1 + q_2 \leqslant q$, $q - q_1 - q_2 \leqslant v$, and since $r = s + q$, we have $m \leqslant s - 1$. Thus $P(x, y)$ contains at most $s - 1$ segments.

Case 2. Suppose x non $\in R_k$ $(k = 1, \ldots, q)$, $y \in R_i$ (i fixed). Let $y \in x_{\alpha-1}x_\alpha$, an edge of R_i. Choose y' in the interior of E_α (see Figure 1). If $E_\alpha \not\subset R_k$ $(k = 1, \ldots, q)$, then by Case 1 $P(x, y')$ has at most $s - 1$ segments. It is easy to see that x and y can be joined by a polygonal arc having at most $s - 1$ segments. Secondly, if $E_\alpha \subset R_j$ (j fixed), then $x_\alpha \in R_i$, $x_\alpha \in R_j$. Let $P(x, x_\alpha)$ and $\sum_{i=1}^{t} \mathfrak{F}_i$ be the quantities in Lemma 2. Since x_α is an improper vertex which is an endpoint of $P(x, x_\alpha)$, and since $P(x, x_\alpha)$ is minimal. there exist at least two faces of T having x_α as a vertex, not belonging to $\sum_{i=1}^{t} \mathfrak{F}_i$, and distinct from F_0 and F'_0 (see Theorem 1). This together with a proof similar to Case 1 implies the following. If x_α is a removable vertex of $P(x, x_\alpha) + x_\alpha y$ or if $x_\alpha y \subset P(x, x_\alpha)$, then $P(x, x_\alpha)$ contains at most $s - 1$ segments. If $P(x, x_\alpha)$ and $x_\alpha y$ are not so related, then $P(x, x_\alpha)$ contains at most $s - 2$ segments. In any case, x and y can be joined by a polygonal arc in T having at most $s - 1$ segments. The same proof holds if x and y are interchanged. If both x and y are contained in the boundaries of pinwheels of $M_r - M_n$, a similar proof applied to x and y simultaneously yields the same conclusions.

THEOREM 7. *Let T be a component of the boundary of a maximal family M_n, and let s be the number of faces of M_n incident with T. Suppose that $s \geqslant 3$, and suppose a slab or half-plane B exists which is incident with T. Then through any point $x \in T$ there passes an infinite polygonal ray in T having at most $s - 2$ segments.*

Proof. If $x \in T \cdot \overline{B}$, then any half-line in $T \cdot \overline{B}$ having x as endpoint will suffice. If $x \in T - T \cdot \overline{B}$, choose a point $y \in T - T \cdot \overline{B}$ which is contained in the interior of an infinite half-line of T. By Theorem 6, there exists a minimal polygonal arc $P(x, y) \subset T$ having at most $s - 1$ segments. If $P(x, y) \cdot \overline{B} = 0$, then B non $\in \mathfrak{F}$, B non $\in \mathfrak{G}$ (see Theorem 1), and it is clear by the arguments

given for Theorem 6, with $v \geqslant 1$, that $P(x, y)$ will contain at most $s - 2$ segments. If $P(x, y) \cdot \overline{B}$ contains an edge of T, then clearly x can be joined to infinity via a portion of $P(x, y)$ and a suitable half-line in $T \cdot \overline{B}$ which together contain at most $s - 2$ segments. If $P(x, y) \cdot \overline{B}$ contains a vertex x' of T which is not incident with an edge of $P(x, y) \cdot \overline{B}$, then x' is improper. Since x' non $\in R_k$ $(k = 1, \ldots, q)$, defined in the proof of Theorem 6, that proof implies that $P(x, y)$ will contain at most $s - 2$ segments. Hence in all cases x can be joined to infinity by an at most $s - 2$ sided polygonal ray in T.

4. C_n-convex sets in the plane. In this section we investigate the complement of an open bounded C_n-convex set.

DEFINITION 8. *A maximal family of disjoint open convex sets M_n is said to be a maximal extension of an open C_n-convex set C_n if $M_n \supset C_n$, and if each member of M_n contains a unique member of C_n.*

THEOREM 8. *The complement of an open bounded C_n-convex set is an L_{n+1} set if $n > 1$. If $n = 1$, the complement is an L_3 set.*

Proof. Let M_n be a maximal extension of C_n, and let M_r be the family defined in Theorem 4, so that $M_r \supset M_n$. Let x_1 and x_2 be any two points in $\overline{M}_r - C_n$, and let K_1 and K_2 be components of M_r such that $x_1 \in \overline{K}_1$, $x_2 \in \overline{K}_2$. The sets K_1 and K_2 need not be distinct. When $n = 1$, the proof is trivial. When $n = 2$, there exist only two components in C_2, so that the boundary of M_2 is a straight line. The proof that x_1 and x_2 can be joined by a polygonal arc L_3 not intersecting C_2 is trivial.

Proof for $n \geqslant 3$. Case 1. Suppose the boundary of M_r has no slabs or half-planes incident with it. In this case the boundary of M_r, denoted by T, must be connected (see Theorem 5). If $x_i \in T$ $(i=1,2)$, relabel it y_i. If x_i non $\in T$, then since each member of C_n is convex, and since each K_i is not a slab or a half-plane there exists a line segment $x_iy_i \subset \overline{M}_r - C_n$ such that $y_i \in T$. By Theorem 6, y_1 and y_2 can be joined by an L_{n-1} polygonal arc in T. Hence, x_1 and x_2 can be joined by an L_{n+1} polygonal arc in $\overline{M}_r - C_n$.

Case 2. Suppose the boundary of T has at least one slab or half-plane incident with it, and suppose that K_1 and K_2 are incident with the same component T_1 of T. Let s denote the number of faces of M_n incident with T_1. If $s = 2$, x_1 and x_2 can be joined by an at most 3-sided polygonal arc in $\overline{M}_r - C_n$. Hence, suppose $s \geqslant 3$. Then either a line passes through x_i not intersecting C_n, or a segment x_iy_i exists such that $y_i \in T_1$, $x_iy_i \cdot C_n = 0$. If both y_1 and y_2 exist, the remainder of the proof is the same as in Case 1. Suppose a line L exists through x_1 not intersecting C_n, and suppose y_2 exists. Then by Theorem 7, a polygonal ray $Q \subset T$ exists through y_2 having at most $s - 2$ segments. Since C_n is bounded, points $z_1 \in L$, $z_2 \in Q$ exist such that $z_1z_2 \cdot C_n = 0$. Hence, it is clear that x_1 and x_2 can be joined by an L_{s+1} $(s \leqslant n)$ polygonal arc in $\overline{M}_r - C_n$. The same proof holds if x_1 and x_2 are interchanged.

Case 3. Suppose T is disconnected, and let K_i be incident with T_i $(i = 1, 2)$, where T_i are components of T with $T_1 \neq T_2$. Let s_i be the number of faces of M_n incident with T_i. Theorem 5 implies $4 \leqslant s_1 + s_2 \leqslant n + 1$. If $s_i = 2$, then x_i can be joined to infinity by a half-line not intersecting C_n. If $s_i \geqslant 3$, then x_i can be joined to infinity by a half-line not intersecting C_n, or a segment x_iy_i exists such thst $y_i \in T_i$, $x_iy_i \cdot C_n = 0$. By applying Theorem 7, then x_i can be joined to infinity in all subcases by a polygonal ray R_i containing at most $s_i - 1$ segments. Since C_n is bounded, there exist points $z_i \in R_i$ such that z_1 and z_2 can be joined by an at most two-sided polygonal arc not intersecting C_n. Hence x_1 and x_2 can be joined by an at most μ-sided polygonal arc in $\overline{M}_r - C_n$, where, by counting, $\mu \leqslant (s_1 - 1) + (s_2 - 1) + 2 = s_1 + s_2 \leqslant n + 1$. This completes the proof.

The expression "C-convex set" was suggested to me by Professor Max Zorn some years ago.

References

[1] Bonnesen and Fenchel, *Theorie der Konvexen Körper* (New York, 1948).

[2] Alfred Horn and F. A. Valentine, *Some properties of L sets in the plane*, Duke Math. J., vol. 16 (1949), 131-140.

[3] M. H. A. Newman, *Elements of the topology of plane sets of points* (Cambridge, 1939).

University of California
at Los Angeles

QUASICONVEX SETS

J. W. GREEN AND W. GUSTIN

Introduction. Let I be the closed real number interval: $0 \leqslant \theta \leqslant 1$. Any subset Δ of I containing at least one number interior to I, will be called a *quasiconvexity generating set.* To each quasiconvexity generating set Δ we associate as follows a generalized notion of convexity, here called quasiconvexity or Δ convexity. Two numbers α and β, one of which belongs to Δ, the other being determined by the relation $\alpha + \beta = 1$, are called complementary ratios of Δ. A set Q in a real vector space is said to be Δ convex if for every pair of complementary ratios α and β in Δ and every pair of points a and b lying in Q the point $\alpha a + \beta b$ also lies in Q.

Quasiconvexity generated by the closed unit interval I evidently coincides with ordinary convexity. We are not, however, interested here in this type of quasiconvexity in that for it our theorems become trivial. More illuminating for our purpose is the quasiconvexity generated by the single self-complementary ratio $\frac{1}{2}$. We shall call this type of quasiconvexity midpoint convexity. It is easily verified that the graph of any solution of the functional equation

$$\varphi(x + y) = \varphi(x) + \varphi(y)$$

in midpoint convex. Such graphs, particularly the discontinuous ones, have been intensively studied and are known to possess many interesting measure and topological properties.

These known properties and other new properties as well follow from our general results on quasiconvex sets.

Notation. We shall denote by X a real normed vector space of finite dimension ν. The norm of a vector x in X will be written $|x|$. Points or vectors in X and real numbers will be denoted by small letters, sets by capital letters.

Set union will be symbolized by $\cup$, set intersection by $\cap$, and set difference by $-$. The symbols $\supset$ and $\subset$ mean "contains" and "is contained in" respectively. The closure of a set E will be denoted by $\overline{E}$, the interior by $\underline{E}$, the boundary by $^{\cdot}E$, and the complement $X - E$ of E in X by CE. The null set is represented by 0.

1. Algebra. Let E be an arbitrary subset of X. The set of all points x in X of the form $x = \alpha a + \beta b$ where a and b lie in E and α and β are complementary ratios of Δ is called the Δ divisor set of E and is denoted by ΔE. Since $x = \alpha x + \beta x$, we see that $E \subset \Delta E$: the divisor operation Δ is ascending. The operation Δ is evidently also increasing in the sense that if $A \subset B$ then $\Delta A \subset \Delta B$.

Received July 1, 1949.

The Δ divisor iterates of E, $\Delta^n E$ ($n = 0, 1, 2, \ldots$), are defined recursively as follows: $\Delta^0 E = E$ and $\Delta^{n+1}E = \Delta\Delta^n E$ for $n \geqslant 0$. Let $\Delta^\omega E$ be the union of all these iterates $\Delta^n E (n \geqslant 0)$; ω may here be regarded in its usual ordinal sense.

A set Q has been defined to be Δ convex if it contains all its Δ divisors: $Q \supset \Delta Q$. Thus Q is Δ convex if and only if $\Delta Q = Q$.

Since the space X is Δ convex, the intersection of any collection of Δ convex sets in X is easily seen to be Δ convex. The intersection of all Δ convex sets in X containing a given set E is then the minimal Δ convex set containing E. It is called the Δ convex hull of E and is denoted by $\Delta[E]$.

THEOREM 1.1. $\Delta[E] = \Delta^\omega E$.

Proof. We first note by induction that $\Delta[E] \supset \Delta^n E$ for $n < \omega$. This is certainly true for $n = 0$, and if true for $n < \omega$ it is also true for $n + 1$, since the set $\Delta[E]$ being Δ convex,

$$\Delta[E] = \Delta\Delta[E] \supset \Delta\Delta^n E = \Delta^{n+1}E.$$

Therefore $\Delta[E] \supset \Delta^\omega E$. On the other hand $\Delta^\omega E$ is a Δ convex set containing E. For let x be a Δ divisor of some two points a and b of $\Delta^\omega E$. Then, since the sets $\Delta^n E$ are ascending, some integer $n < \omega$ exists such that $a, b \subset \Delta^n E$. Therefore

$$x \subset \Delta(a, b) \subset \Delta\Delta^n E = \Delta^{n+1}E \subset \Delta^\omega E,$$

whence $\Delta^\omega E$ is Δ convex, so that $\Delta[E] \subset \Delta^\omega E$. This completes the proof.

The set $\Delta^* = \Delta[0, 1]$ is evidently a quasiconvexity generating set, and is, moreover, Δ convex. From the linear character of the space X we see that $\Delta^*(a, b) = \Delta[a, b]$.

This set Δ^* plays a special role in the theory of Δ convexity. It is particularly important in the discussion of what we shall call equivalent quasiconvexity generating sets. Let $\{\Delta\}$ denote the class of all Δ convex sets. We say that Δ generates $\{\Delta\}$. Two quasiconvexity generating sets Δ_1 and Δ_2 will be called equivalent, and we write $\Delta_1 \sim \Delta_2$, if they generate the same sets; that is, if $\{\Delta_1\} = \{\Delta_2\}$. Clearly $\sim$ is a true equivalence relation.

THEOREM 1.2. $\Delta^* \sim \Delta$; $\Delta^*[E] = \Delta[E]$; $\Delta^{**} = \Delta^*$.

Proof. Since $\Delta^* \supset \Delta$, every Δ^* convex set is evidently Δ convex. On the other hand every Δ convex set Q is also Δ^* convex. For let x be a Δ^* divisor of some two points a and b of Q; then

$$x \subset \Delta^*(a, b) = \Delta[a, b] \subset \Delta[Q] = Q.$$

Therefore $\Delta^* \sim \Delta$. Since the Δ^* convex set $\Delta^*[E]$ is Δ convex, $\Delta^*[E] \supset \Delta[E]$; and since the Δ convex set $\Delta[E]$ is Δ^* convex, $\Delta[E] \supset \Delta^*[E]$. Consequently $\Delta^*[E] = \Delta[E]$. Finally we have

$$\Delta^{**} = \Delta^*[0, 1] = \Delta[0, 1] = \Delta^*.$$

THEOREM 1.3. $\{\Delta_1\} \subset \{\Delta_2\}$ *if and only if* $\Delta_1^* \supset \Delta_2^*$.

Proof. If $\Delta_1^* \supset \Delta_2^*$, then every Δ_1^* convex set is plainly Δ_2^* convex, and hence every Δ_1 convex set is Δ_2 convex. On the other hand if every Δ_1 convex set is Δ_2 convex, then, since Δ_1^* is Δ_1 convex and consequently Δ_2 convex, we have

$$\Delta_1^* = \Delta_2[\Delta_1^*] \supset \Delta_2[0, 1] = \Delta_2^*.$$

It follows from this result that $\Delta_1 \sim \Delta_2$ if and only if $\Delta_1^* = \Delta_2^*$. Thus the set Δ^* is the maximal quasiconvexity generating set equivalent to Δ. The equivalence relation $\sim$ divides all the quasiconvexity generating sets, essentially all subsets of I, into pairwise disjoint non-null equivalence classes each of which may be uniquely represented by its maximal element namely by Δ^*, where Δ is any element in the class.

Let Q be a given Δ convex set and α and β be positive complementary ratios of Δ^*. We shall in the sequel make frequent use of the following three types of projection mappings, which since the ratios α and β are chosen from Δ^* will be called Δ^* projections.

The projection f defined by the equation $f(x) = \alpha s + \beta x$ is a contraction toward the point s. If $s \subset Q$, then $f(Q) \subset Q$; for if $s \subset Q$ and $x \subset Q$, then $f(x) \subset \Delta^* Q = Q$.

The projection f defined by the equation $x = \alpha s + \beta f(x)$ is an expansion away from the point s. If $s \subset Q$, then $f(CQ) \subset CQ$; for if $s \subset Q$ and $f(x) \subset Q$, then $x \subset \Delta^* Q = Q$.

The projection f defined by the equation $s = \alpha x + \beta f(x)$ is a reflection through the point s. If $s \subset CQ$, then $f(Q) \subset CQ$; for if $x \subset Q$ and $f(x) \subset Q$, then $s \subset \Delta^* Q = Q$.

In each of the projections f defined above the point s is the centre of projection and any image set is similar to its original. The projection f is a topological mapping, and its inverse f^{-1} is also a projection: if f is a Δ^* contraction, expansion, or reflection with centre s, then f^{-1} is a Δ^* expansion, contraction, or reflection respectively also with centre s.

We may summarize the above results on Δ^* projections as follows: a Δ^* contraction of Q toward a point of Q lies in Q; a Δ^* expansion of CQ away from a point of Q lies in CQ; a Δ^* reflection of Q through a point of CQ lies in CQ.

2. Density. In this section we investigate some elementary topological properties of quasiconvex sets. All the results here stem from the following density property of Δ^*.

THEOREM 2.1. Δ^* *is dense in* I.

Proof. Suppose to the contrary that the open set $I - \overline{\Delta^*}$ is non-null. Let J be an open interval component of this set with end points a and b, which lie in $\overline{\Delta^*}$. Thus there exist point sequences a_n and b_n of Δ^* with $a_n \to a$ and $b_n \to b$. Let α and β be positive complementary ratios of Δ. Then the point $x = \alpha a + \beta b$

lies in J, and the points $x_\varepsilon = \alpha a_\varepsilon + \beta b_\varepsilon$ lie in $\Delta\Delta^* = \Delta^*$. Now $x_\varepsilon \to x$; whence x lies in $\overline{\Delta^*}$ and hence not in J. This contradiction proves the theorem.

THEOREM 2.2. *A quasiconvex set is dense in its convex hull.*

Proof. Every point in the convex hull of a set lies in the convex hull of some finite subset of that set. Thus it suffices to prove that $\Delta[A]$ is dense in $I[A]$ for every finite set A. This is clearly true for the null set and hence by induction is true for every finite set if it is true for a finite set A containing a point a whenever it is true for the set $A - a$. To demonstrate this let x be a point of $I[A]$; then x may be expressed in the form $x = \alpha a + \beta b$ where α and β are complementary ratios of I and $b \subset I[A - a]$. Since Δ^* is dense in I there exist complementary ratios α_ε and β_ε of Δ^* with $\alpha_\varepsilon \to \alpha$ and $\beta_\varepsilon \to \beta$. Furthermore by the induction hypothesis points b_ε of $\Delta[A - a]$ exist with $b_\varepsilon \to b$. Thus the points $x_\varepsilon = \alpha_\varepsilon a + \beta_\varepsilon b_\varepsilon$ lie in $\Delta^*[A] = \Delta[A]$. But clearly $x_\varepsilon \to x$, so $\Delta[A]$ is dense in $I[A]$.

The interior of the convex hull of a set E will be called the near interior of E and the complement of E in its near interior the near complement of E. Thus $\underline{I(E)}$ is the near interior of E and $\underline{I(E)} - E$ the near complement of E. Note that E is non-planar if and only if its near interior is non-null. We shall say of a non-planar set E that it is nearly convex if it contains its near interior: $E \supset \underline{I(E)}$, that is, if its near complement is null.

Nearly convex sets play an important role in the theory of quasiconvexity. Thus many of our theorems read: A quasiconvex set having such and such a property is nearly convex. We assume that any quasiconvex set forming the subject of a theorem is non-planar, so that the notion of near convexity is applicable to it. The hypotheses of the theorem usually ensure this.

LEMMA 2.3. *Let Q be a Δ convex set with near interior G; let S_q be an open sphere about $q \subset Q$; and let p be a point of G different from q. Then an open sphere $S_p \subset G$ about p and positive complementary ratios α, $\beta \subset \Delta^*$ exist such that for every non-null open subset V of S_p a point $a \subset Q$ can be found for which the expansion f away from a defined by the equation $x = \alpha a + \beta f(x)$ has the property that $q \subset f(V) \subset S_q$.*

Proof. Let p be the origin and let the radius of S_q be $2\epsilon\rho$ where $\rho = |q| > 0$. We may assume $\epsilon \leqslant 1$. Since $p \subset G$, some sphere S, say of radius $2\lambda\rho$, about p lies in G. Let α and β be positive complementary ratios of Δ^* chosen so that $\beta < \lambda/(1 + \lambda)$ whence $\beta/\alpha < \lambda$. Let S_p be the open sphere about p of radius $\epsilon\beta\rho < \lambda\rho$. Thus $S_p \subset S \subset G$. Consider the expansion g away from q defined by the equation $x = \alpha g(x) + \beta q$. Evidently for $x \subset S_p$ we have

$$| g(x) | = \frac{1}{\alpha} | x - q | < \frac{1}{\alpha} (\epsilon\beta\rho + \beta\rho) < 2\lambda\rho,$$

whence $g(V) \subset g(S_p) \subset S \subset G$ for any open subset V of S_p. Since Q is dense in S and hence in the open set $g(V)$, some point $a \subset Q \cap g(V)$ can be selected.

Let $v = g^{-1}(a)$; then $v \subset V$ and $g(v) = a \subset Q$. Thus $v = \alpha a + \beta q$. Now consider the expansion f away from a defined by the equation $x = \alpha a + \beta f(x)$. We observe that $q = f(v)$, so for $x \subset S_p$

$$|f(x) - q| = \frac{1}{\beta}|x - v| < \frac{1}{\beta}(\epsilon\beta\rho + \epsilon\beta\rho) = 2\epsilon\rho$$

whence $f(S_p) \subset S_q$. Thus we conclude that $q = f(v) \subset f(V) \subset f(S_p) \subset S_q$.

THEOREM 2.4. *A quasiconvex set with non-null interior is nearly convex.*

Proof. Let Q be a Δ convex set containing an open sphere S_q with center q. We are to show that Q contains its near interior G. Since $q \subset Q$, we consider to this end any point p in G different from q. According to the preceding lemma there exists a point $a \subset Q$ and a Δ^* expansion f away from a with the property that $f(p) \subset S_q \subset Q$ whence $p \subset f^{-1}(Q)$. Now f^{-1} is a Δ^* contraction toward the point $a \subset Q$, so $p \subset f^{-1}(Q) \subset Q$. Therefore $G \subset Q$.

3. Measure. Let μ^* be an outer measure function and μ_* the corresponding inner measure function defined on subsets of X. If E is a measurable set then $\mu^*(E) = \mu_*(E)$ and we write $\mu(E)$ for this common value. We assume that μ^* and μ_* are homogeneous measures in the following sense: if f is a projection with ratio of similarity θ, then for every set E we have $\mu^*(f(E)) = \theta^r\mu^*(E)$ and $\mu_*(f(E)) = \theta^r\mu_*(E)$.

THEOREM 3.1. *The near complement of a quasiconvex set of positive outer measure has zero inner measure.*

Proof. Let Q be a Δ convex set of positive outer measure. Suppose, contrary to the theorem, that the near complement P of Q has positive inner measure. Let p be a point of inner density of P and let $\eta = \frac{1}{2}$. Then an open sphere about p of radius ρ exists such that for every smaller concentric open sphere S_p we have

$$\mu_*(S_p \cap P) > \eta\mu(S_p).$$

Now let q be a point of outer density of Q. Then an open sphere S_q of radius $\rho_q < \rho$ exists with

$$\mu^*(S_q \cap Q) > (1 - \eta^{r+1})\,\mu(S_q),$$

whence

$$\mu_*(S_q - Q) < \eta^{r+1}\mu(S_q).$$

According to Lemma 2.3 a point of Q and a Δ^* expansion f away from this point can be found with the property that $|b - q| < \eta\rho_q$, where $b = f(p)$. Let S_b be the sphere about b of radius $\rho_b = \eta\rho_q$. Then $\mu(S_b) = \eta^r\mu(S_q)$, and, since $\eta = \frac{1}{2}$, $S_b \subset S_q$. Furthermore, the inverse set $S_p = f^{-1}(S_b)$, being a contraction of S_b, is an open sphere about $p = f^{-1}(b)$ with radius $\rho_p < \rho_b < \rho_q < \rho$. Hence $\mu_*(S_p \cap P) > \eta\mu(S_p)$ and consequently

$$\mu_*(f(S_p \cap P)) > \eta\mu(f(S_p)) = \eta\mu(S_b) = \eta^{r+1}\mu(S_p).$$

Since f is a Δ^* expansion away from a point of Q, we have $f(P) \subset f(CQ) \subset CQ$. Therefore

$$f(S_p \cap P) = f(S_p) \cap f(P) \subset S_b \cap CQ \subset S_q - Q$$

so

$$\mu_*(f(P \cap S_p)) \leqslant \mu_*(S_q - Q) < \eta^{r+1}\mu(S_q)$$

in contradiction to a preceding inequality. This contradiction proves that P has zero inner measure.

Let P be the near complement of a quasiconvex set Q. Under the assumption that Q has positive outer measure we have shown that P has zero inner measure. Under the stronger hypothesis that Q has positive inner measure we now prove the stronger conclusion that P is the null set, that is, Q is nearly convex.

THEOREM 3.2. *A quasiconvex set of positive inner measure is nearly convex.*

Proof. Let Q be a Δ convex set of positive inner measure. Then Q contains a measurable set F of positive measure. Let q be a point of density of F and let $\eta = \frac{1}{2}$. Then there exists an open sphere S about q of radius ρ such that

$$\mu(F \cap S) > (1 - \alpha_r \eta^r)\mu(S)$$

where $\alpha_r = \alpha^r/(\alpha^r + \beta^r)$, α and β being positive complementary ratios of Δ^* with $\alpha \leqslant \beta$. Let S_q be the sphere about q of radius $\eta\rho$. We contend that $S_q \subset Q$. Suppose, to the contrary, that some point p, which we may assume to be the origin of S_q, does not lie in Q. Let S_p be the sphere about p of radius $\eta\rho$; then, since $\eta = \frac{1}{2}$, $S_p \subset S$. Let $F_p = F \cap S_p$; then

$$\mu(F_p) = \mu(F \cap S_p) = \mu(F \cap S) - \mu(F \cap S - S_p).$$

Consequently

$$\mu(F \cap S - S_p) \leqslant \mu(S - S_p) = \mu(S) - \mu(S_p) = (1 - \eta^r)\mu(S),$$

wherefore

$$\mu(F_p) > [(1 - \alpha_r \eta^r) - (1 - \eta^r)]\mu(S) = \beta_r \mu(S_p),$$

the ratio β_r being complementary to α_r. Let f be the Δ^* reflection through the point p of CQ defined by the equation $p = \alpha x + \beta f(x)$. Therefore $f(F_p) \subset f(Q) \subset CQ \subset CF$. Moreover, since $\alpha \leqslant \beta$, we have

$$|f(x)| = \frac{\alpha}{\beta}|x| \leqslant |x|$$

whence $f(F_p) \subset S_p$. Consequently the set F_p and its reflection $f(F_p)$ are disjoint measurable subsets of S_p, so that

$$\mu(S_p) \geqslant \mu(F_p) + \mu(f(F_p)) = \left(1 + \frac{\alpha^r}{\beta^r}\right)\mu(F_p) = \frac{1}{\beta_r}\mu(F_p)$$

in contradiction to a preceding inequality. This contradiction proves that the quasiconvex set Q contains the sphere S_p and hence is nearly convex.

THEOREM 3.3. *Quasiconvexity generated by a set of positive inner measure is equivalent to convexity.*

Proof. Let Δ be a quasiconvexity generating set of positive inner measure. Consider the Δ convex set $\Delta^* \supset \Delta$, and note that $\Delta \sim \Delta^* = I$.

We remark that quasiconvexity generated by a set of measure zero may be equivalent to convexity. The Cantor ternary set is an example of such a set.

Our theorems on measure of quasiconvex sets may be stated as follows: Every quasiconvex set Q is either extremely measurable or extremely non-measurable. By this we mean that if Q is measurable its measure is as small or as large as possible, and if Q is non-measurable its inner measure is as small as possible and its outer measure as large as possible: zero being as small a measure as possible and the measure of the convex hull of Q being as large a measure as possible.

4. Subcontinua. In this section we investigate what happens when a quasiconvex set or its near complement contains a certain type of continuum. We show that a quasiconvex set containing a non-planar continuum is nearly convex and that a quasiconvex set whose near complement contains a certain type of non-planar continuum is zero dimensional in the topological sense of dimension.

THEOREM 4.1. *A quasiconvex set containing a non-planar continuum is nearly convex.*

Let Q be a Δ convex set containing a non-planar continuum K. We may suppose that K contains the origin and ν linearly independent vectors k_λ $(\lambda = 1, \ldots, \nu)$. Let α and β be positive complementary ratios of Δ. If $x_1, \ldots x_\nu$ are points of Q then it is easily verified by taking successive α, β linear combinations that the point

$$x = \theta_1 x_1 + \ldots + \theta_\nu x_\nu$$

also lies in Q where $\theta_1 = \alpha^{\nu-1}$ and $\theta_\lambda = \alpha^{\nu-\lambda}\beta$ for $\lambda = 2, \ldots, \nu$. Let $H_\lambda (\lambda = 1, \ldots \nu)$ be the θ_λ contraction of the continuum K toward the origin. Thus H consists of all points of the form $\theta_\lambda x$ for $x \subset K$. We have just indicated that the vector sum H of the ν continua H_λ lies in Q. However, since the vectors $h_\lambda = \theta_\lambda k_\lambda$ form a basis for X, this vector sum set H has a non-null interior. (For the proof of this see the paper which follows entitled, *On the vector sum of continua.*) Therefore Q has a non-null interior and hence is nearly convex.

We now digress into some lemmas concerning convex sets.

Let K be a compact convex set in X. We shall call a point a an apex of K if for every neighbourhood N of a there exists an open space H containing a whose intersection with K lies in $N : H \cap K \subset N$. An apex of K is evidently

a boundary point of K, but not necessarily conversely. Let $A(K)$ be the set of apices of K.

LEMMA 4.2. *$A(T \cap K) = T \cap A(K)$ for every supporting plane T of K; if $I[E] = K$, then $E \supset A(K)$; $I[A(K)] = K$.*

Proof. It is evident that $T \cap A(K) \subset A(T \cap K)$. Suppose then that $a \subset A(T \cap K)$. Thus for every neighbourhood N of a there exists a half plane V of dimension $\nu - 1$ open in T such that $a \subset V \cap (T \cap K) = V \cap K \subset N$. Let L be that linear subspace of dimension $\nu - 2$, a plane in T, which bounds V; and let T' be a variable plane of dimension $\nu - 1$ which contains L and is different from T. Furthermore let V' be that half plane open in T which is bounded by $L = T' \cap T$ and lies on the same side of the supporting plane T of K as does K; and let H' be that open half space of dimension ν bounded by T' which contains V. Then for some H' we have $H' \cap K \subset N$; else $H' \cap K - N \neq 0$ for all H', so that by choosing T' approaching T in such a way that V' approaches V it would follow from the compactness of $K - N$ that $V \cap K - N \neq 0$. This completes the proof of the first part of the lemma.

To prove the second part, suppose that $I[E] = K$ but that $a \subset A(K) - E$. Since $a \subset A(K) \subset K = I[E]$ there exists a finite set $F \subset E \subset K$ whose convex hull contains a although the set F itself does not contain a. Therefore the open set CF is a neighbourhood of the apex a of K, so an open half space H containing a exists such that $H \cap K \subset CF$, whence $H \cap F = H \cap F \cap K = 0$. Therefore $F \subset CH$; that is, the closed half space CH is a convex set containing F but not a, in contradiction to $a \subset I[F]$.

The proof of the third part of the lemma proceeds by induction on the dimension of K. It is clearly true for dimension 1. Assume it true for dimension $\nu - 1 \geqslant 1$, and let K be of dimension ν. Furthermore let t be any boundary point of K and let T be a supporting plane to K at t. From the induction hypothesis and the first part of the lemma we see that

$$t \subset T \cap K = I[A(T \cap K)] = I[T \cap A(K)] \subset I[A(K)].$$

Thus every boundary point of K belongs to the convex set $I[A(K)]$ so that $K \subset I[A(K)]$. It is obvious that $K \supset I[A(K)]$. This concludes the proof of the lemma. We note that $A(K)$ is the minimal set whose convex hull is K.

We shall say that a set E is indented if E together with some plane T bounds a non-null bounded open set $W : {}^{\cdot}W \subset T \cup E$. A point $p \subset E$ will be called an indentation point of E provided every neighbourhood of p contains an indented subset of E.

LEMMA 4.3. *Every indented set contains an indentation point.*

Proof. Let E be an indented set; and let T be a plane and W a non-null bounded open set such that ${}^{\cdot}W \subset T \cup E$. Consider the compact set $K = I[W]$. Since $K = I[A(K)]$ is non-planar, the set $A(K)$ is also non-planar. Therefore $A(K)$ contains some point, say p, not in the plane T. Now $A(K)$ is the minimal

set whose convex hull is K, so $p \subset \overline{W}$. Moreover, the apex p is not interior to $\overline{W}$; hence $p \subset {}^{\cdot}W \subset T \cup E$. But $p \not\subset T$, so $p \subset E$. Consider any neighbourhood N_p of p. Since $p \not\subset T$, a neighbourhood N of p can be found such that $\overline{N} \cap T = 0$ and $\overline{N} \subset N_p$. Now p is an apex of K, so there exists an open half space H containing p such that $H \cap K \subset N$. Since $p \subset H \cap {}^{\cdot}W$, the set $V = H \cap W$ is a non-null open set containing p on its boundary ${}^{\cdot}V$. Evidently $V \subset N$, so ${}^{\cdot}V \cap T \subset \overline{N} \cap T = 0$. Consequently we see from the inclusion ${}^{\cdot}V \subset {}^{\cdot}H \cup {}^{\cdot}W \subset {}^{\cdot}H \cup T \cup E$ that ${}^{\cdot}V \subset {}^{\cdot}H \cup E$, ${}^{\cdot}H$ being the bounding plane of H. Thus ${}^{\cdot}V$ is an indented subset of E and ${}^{\cdot}V \subset \overline{N} \subset N_p$. Since this is so for any neighbourhood N_p of p, we conclude that the point $p \subset E$ is an indentation point of E.

LEMMA 4.4. *Every neighbourhood of an indentation point of the near complement P of a quasiconvex set contains a non-null open subset with boundary in P.*

Proof. Let Q be a Δ convex set with near interior G and near complement P; and let N be a neighbourhood of an indentation point p of P. Thus the set $N \cap G$ is a neighbourhood of p and hence contains some open sphere S_p about p. Let the radius of S_p be 5ρ and let S be the open sphere about p of radius ρ. Since p is an indentation point of P, a plane T and a non-null bounded open set W exist such that ${}^{\cdot}W \subset T \cup P$ and $W \subset S$. We shall for convenience assume that the plane T contains the origin. Let φ be a linear functional vanishing on T such that W intersects the open half space $\varphi > 0$. Define 4μ to be the upper bound of $\varphi(w)$ as w ranges over the non-null bounded set W; then $\mu > 0$. Let H be the open half space $\varphi > 8\mu$; and let K be the closed half space $\varphi \leqslant 0$. Consider the expansion f_q away from q defined by the equation $x = \alpha q + \beta f_q(x)$ where α and β are fixed positive complementary ratios of Δ^* so chosen that $\frac{1}{3} < \alpha < \frac{3}{8}$. We assert that the open expansion sets $f_q(W)$ cover $K \cap \overline{W}$ as q ranges over $Q \cap H$. To prove this let k be a given point of $K \cap \overline{W}$; and let g be the expansion away from k defined by the equation $x = \alpha g(x) + \beta k$. By definition of μ some point $w \subset W$ exists such that $\varphi(w) > 3\mu$. Since $\varphi(k) \leqslant 0$ and $\alpha < \frac{3}{8}$ we have

$$\varphi(g(w)) = \frac{1}{\alpha}\varphi(w) - \frac{\beta}{\alpha}\varphi(k) > 8\mu,$$

so that $H \cap g(W) \neq 0$. Furthermore, for any point $w \subset W \subset S$ we have

$$g(w) - p = \frac{1}{\alpha}(w - p) - \frac{\beta}{\alpha}(k - p),$$

so that

$$| g(w) - p | \leqslant \frac{\rho}{\alpha} + \frac{\beta\rho}{\alpha} < 5\rho$$

since $\alpha > \frac{1}{3}$ and $k \subset \overline{W} \subset \overline{S}$. Therefore $g(W) \subset S_p \subset G$. The set $H \cap g(W)$ is then a non-null open subset of G. Since Q is dense in G, some point $q \subset Q \cap$

$H \cap g(W)$ exists. Let $w = g^{-1}(q)$. Then $w = \alpha q + \beta k \subset W$, so that $k = f_q(w) \subset f_q(W)$.

This proves that the open sets $f_q(W)$ cover the compact set $K \cap \overline{W}$ as q ranges over $Q \cap H$. Consequently a finite subset Y of $Q \cap H$ exists such that the sets $f_y(W)$ cover $K \cap \overline{W}$ as y ranges over Y. Define $U = \bigcup_y f_y(W)$. Thus U is an open set and $K \cap \overline{W} \subset U$. Now $f_y(^{\cdot}W) \subset K \cup CQ$. For since f_y is a Δ^* expansion away from a point of Q we have $f_y(CQ) \subset CQ$. And since $\varphi(y) > 0$ and $\varphi(k) \leqslant 0$, we have

$$\varphi(f_y(k)) = \beta^{-1}[\varphi(k) - \alpha\varphi(y)] < 0,$$

whence $f_y(K) \subset K$. But $^{\cdot}W \subset K \cup CQ$, so

$$f_y(^{\cdot}W) \subset f_y(K \cup CQ) = f_y(K) \cup f_y(CQ) \subset K \cup CQ.$$

Therefore $^{\cdot}U \subset \bigcup_y f_y(^{\cdot}W) \subset K \cup CQ$, the union being finite.

Consider the following open subset of W: $V = W - \overline{U}$. We shall show that V is a non-null open subset of N whose boundary lies in P. Evidently V is an open subset of $W \subset N$. We note that $^{\cdot}V \subset CV = CW \cup \overline{U}$. Since U and V are disjoint open sets, U and $\overline{V}$ are also disjoint, so $^{\cdot}V \subset \overline{V} \subset CU$. It is clear that $^{\cdot}V \subset \overline{V} \subset \overline{W}$. Combining these inclusions with the inclusions $^{\cdot}W \subset K \cup CQ$ and $^{\cdot}U \subset K \cup CQ$, we obtain the result that $^{\cdot}V \subset {}^{\cdot}W \cup {}^{\cdot}U \subset K \cup CQ$. But since $^{\cdot}V \cap K \subset \overline{W} \cap K \subset U$ and $^{\cdot}V \subset CU$ and $^{\cdot}V \subset G$, we conclude that $^{\cdot}V \subset G \cap CQ = P$.

To complete the demonstration we must show that V is non-null. To do this we prove that each of the open sets $f_y(W)$ composing U lies in the open half space $\varphi < 2\mu$. For let $y \subset Y$ and $w \subset W$; then $\varphi(y) > 8\mu$ and $\varphi(w) < 4\mu$, so that

$$\varphi(f_y(w)) = \frac{1}{\beta}\varphi(w) - \frac{\alpha}{\beta}\varphi(y) < \frac{4\mu}{\beta} - \frac{8\alpha\mu}{\beta} < 2\mu$$

since $\alpha > \frac{1}{3}$. The upper bound of $\varphi(w)$ for $w \subset W$ is 4μ, so we see that $V = W - \overline{U} \neq 0$. This completes the proof of the lemma.

THEOREM 4.5. *A quasiconvex set whose near complement is indented has topological dimension zero.*

Proof. Let Q be a Δ convex set whose near complement P is indented; and let p be a fixed indentation point of P. Consider any point $q \subset Q$ and any open sphere S_q about q. Since p lies in the near complement G of Q there exists according to Lemma 2.3 a sphere $S_p \subset G$ about p and positive complementary ratios $\alpha, \beta \subset \Delta^*$ such that for any non-null open subset V of G a point $a \subset Q$ can be found for which the Δ^* expansion away from a defined by the equation $x = \alpha a + \beta f(x)$ has the property that $q \subset f(V) \subset S_q$. Let V be a non-null open subset of S_p, such as constructed in the preceding lemma, with boundary in CQ. Therefore the neighbourhood $f(V)$ of Q lies in S_q and its boundary, being a Δ^* expansion away from a point of Q of the set $^{\cdot}V \subset CQ$. This is so for every

point q of Q and every sphere S_q about q. Thus we conclude that Q has topological dimension zero.

For $\nu = 2$ this theorem takes the following form.

THEOREM 4.6. *In a two dimensional vector space any quasiconvex set whose near complement contains a non-linear closed connected set has topological dimension zero.*

Proof. The proof consists in showing that any non-linear closed connected set F is indented. Now either F is convex and the result is obvious, or else there exists a line L whose intersection with F is not connected. Therefore, according to a theorem of Janiszewski, one of the components of $F \cup L$ must be bounded; so F is indented.

5. Connectedness. Many examples of pathological connected sets may be constructed in a more or less systematic fashion as graphs of solutions of the functional equation $\varphi(x + y) = \varphi(x) + \varphi(y)$ [17]. It is thus of interest to investigate the connectedness of quasiconvex sets and their near complements.

LEMMA 5.1. *If E is connected, then ΔE is connected.*

Proof. Let p be any point of ΔE. Then p may be expressed in the form $p = \alpha a + \beta b$ where α, β are complementary ratios of Δ and $a, b \subset E$. Consider the contraction f defined by the vector formula $f(x) = \alpha a + \beta x$. The set $f(E)$, being a Δ contraction of E toward a point of E, lies in ΔE. Furthermore, $f(E)$ is similar to E and hence is connected. Now $p = f(b)$ lies in $f(E)$ and $a = f(a)$ also lies in $f(E)$. Therefore $f(E)$ is a connected set containing p and intersecting the connected set E. This proves ΔE connected.

THEOREM 5.2. *A quasiconvex set containing a non-planar connected set is connected.*

Proof. Let Q be a Δ convex set containing a non-planar connected set E and let Q' be that component of Q which contains E. Then Q' is closed in Q and $\Delta Q' \subset \Delta Q = Q$. According to the preceding lemma $\Delta Q'$ is a connected superset of Q'. Consequently $\Delta Q' = Q'$, so Q' is a non-planar connected Δ convex set. It is therefore dense in some open set namely its near interior G'. We assert that $Q' = Q$. For suppose to the contrary that Q' is a proper subset of Q. Then since Q' is closed in Q there exists a point q of Q at a positive distance from Q. Evidently a slight Δ^* contraction f toward q can be obtained such that the convex open sets G' and $f(G')$ interesct. Now Q' is a connected subset of Q dense in G', and $f(Q')$ is a connected subset of Q, dense in $f(G')$; the union $Q' \cup f(Q')$ is then connected. But Q' is a component of Q, so $f(Q') \subset Q'$. On the other hand the contraction set $f(Q')$ lies slightly closer to q than does Q'. This contradiction proves $Q' = Q$, wherefore Q is connected.

Let A be a plane or planar portion. We shall say that a set E is semiconnected parallel to A if no plane parallel to A separates E. The set E will be called semiconnected if it is semiconnected to every plane.

THEOREM 5.3. *A quasiconvex set containing a planar portion and semiconnected parallel to this planar portion is connected.*

Proof. Let Q be a Δ convex set containing a planar portion A and semiconnected parallel to A. Suppose, contrary to the theorem, that Q is not connected. Then Q is separated by some closed set F cutting X. Thus F also cuts the near interior G of Q. Let V be a component of $G - F$, and W a component of $G - \overline{V}$. The space X is locally connected, so the sets W and V, being components of open sets, are open subsets of G. Since any point in $G - F$ lies in some open component of $G-F$, we see that $G \cap^{\cdot} V \subset F$. Similarly $G \cap^{\cdot} W \subset \overline{V}$. Now $W \cap \overline{V} = 0$ so $\overline{V} \cap W = 0$; whence $\overline{V} \subset CW$ and $\overline{W} \subset CV$. It is clear from this that the boundary $B = G \cap^{\cdot} W$ of W in G is given by

$$B = G \cap \overline{W} \cap \overline{V} = G \cap^{\cdot} W \cap^{\cdot} V \subset F \subset CQ.$$

Furthermore, since G is connected and V and W are non-null, this set B also is non-null.

Now consider a point $p \subset B$. Evidently a point $a \subset Q$ and an open sphere S about p and lying in G can be found such that any Δ^* contraction of the planar portion $A \subset Q$ toward the point $a \subset Q$ which intersects S also cuts S. We shall call the intersection sets with S of these Δ^* contractions Q-discs. Thus the Q-discs are parallel planar portions dense in S which cut S and lie in Q. Hence they do not intersect B, so any Q-disc which intersects W or V lies wholly in W or V. Since p is a limit point of both W and V, it is also a limit point of discs lying in W and of discs lying in V. Let D_p be the disc formed by intersection with S of the plane through p parallel to A. Therefore $D_p \subset \overline{W}$ and $D_p \subset \overline{V}$. Also $D_p \subset S \subset G$. Consequently $D_p \subset G \cap \overline{W} \cap \overline{V} = B$. Thus every point $p \subset B$ lies in some disc D_p open in the plane parallel to A through p.

Let D be the intersection with G of a plane parallel to A which intersects the non-null closed set B. Then D is convex and hence connected. Moreover, $B \cap D$ is a non-null set closed in D, which, as we have just shown, is also open in D. Thus $B \cap D = D$, so that $D \subset B \subset F$. Therefore $\overline{D} \subset F \subset CQ$, so the plane extension T of D does not intersect Q. This plane T, parallel to A, then separates Q in contradiction to the semiconnectedness of Q parallel to A. This contradiction proves Q connected.

The following theorem is similar to the theorem just proved and can be proved in a similar fashion: A quasiconvex set containing a linear portion A is connected if it cannot be separated by a cylinder parallel to A.

THEOREM 5.4. *A bounded quasiconvex set containing a planar portion and semiconnected parallel to this planar portion is nearly convex.*

Proof. Let Q be a bounded Δ convex set containing a planar portion A and semiconnected parallel to A. We may assume that A lies in the near interior G of Q, for we could otherwise replace A by a suitable Δ contraction of A which does lie in G. Moreover, we shall for convenience suppose that A contains the origin. If T is the plane containing A, then a radius $\rho > 0$ exists such that

every point of T in the ρ sphere about the origin lies in A. Let φ be a linear functional of norm 1 vanishing on T. Since Q is semiconnected parallel to T, the set $\varphi(Q)$ of real numbers contains an open λ neighbourhood of 0 for some $\lambda > 0$. Let $\epsilon = \min(\lambda, \rho)$ and let $\mu\epsilon$ be a bound of the bounded set Q. Choose complementary ratios α and β of Δ^* such that $0 < \alpha < (1+\mu)^{-1}$; and define $\eta = \min[\alpha, 1-\alpha(1+\mu)]$. We contend that Q contains the open sphere S of radius $\eta\epsilon$ about the origin. To prove this consider a point $x \subset S$; thus $|x| < \eta\epsilon$. Now

$$\alpha^{-1}\varphi(x) \leqslant \alpha^{-1}|x| < \alpha^{-1}\eta\epsilon \leqslant \epsilon \leqslant \lambda;$$

so there exists a point $q \subset Q$ such that $\varphi(x) = \alpha\varphi(q)$. Consider the point t defined by the equation $x = \alpha q + \beta t$. We see that

$$\varphi(t) = \beta^{-1}[\varphi(x) - \alpha\varphi(q)] = 0;$$

so $t \subset T$. Also by choice of η we have

$$|t| \leqslant \beta^{-1}(|x| + \alpha|q|) < \beta^{-1}(\eta\epsilon + \alpha\mu\epsilon) \leqslant \epsilon \leqslant \rho,$$

whence $t \subset Q$. Thus $x = \alpha q + \beta t \subset \Delta^* Q = Q$. This proves that Q contains the open sphere S and hence is nearly convex.

A result similar to the theorem just proved can be similarly proved, namely: A bounded quasiconvex set Q is nearly convex if it contains a linear portion A and if every line parallel to this linear portion intersecting the near interior of Q also intersects Q, that is, if Q is opaque parallel to A.

THEOREM 5.5. *If the near complement of a quasiconvex set is semiconnected it is also connected.*

Proof. Let Q be a quasiconvex set with near interior G whose near complement P is semiconnected. Suppose, contrary to the theorem, that P is not connected. Then some set F in $\dot{G}$ cuts G and separates P. Let V be a component of $G - F$ and W a component of $G - \overline{V}$. Then as in 5.3 the boundary B of W in G is non-null and $B \subset F \subset Q$. Let $G_\kappa (\kappa = 1, 2, \ldots)$ be a sequence of bounded non-null convex open sets intersecting B, the sequence strictly increasing to G in the sense that $\overline{G}_\kappa \subset G_{\kappa+1}$ and $G = \cup G_\kappa$. We define sets W_κ $(\kappa = 0, 1, 2, \ldots)$ recursively as follows. Let $G_0 = 0$ and $W_0 = 0$, and suppose W_κ to be a connected open subset of $G_\kappa - B$. Since B cuts G and intersects $G_{\kappa+1}$ we see that W and V also intersect $G_{\kappa+1}$ so that B cuts $G_{\kappa+1}$. Therefore the set $G_{\kappa+1} - B$ contains the connected open set W_κ and possesses at least two components. Let $V_{\kappa+1}$ be a component such that $V_{\kappa+1} \cap W_\kappa = 0$ and let $W_{\kappa+1}$ be the component of $G_{\kappa+1} - \overline{V}_{\kappa+1}$ containing W_κ. Thus $W_\kappa \subset W_{\kappa+1}$. Since $G_\kappa (\kappa = 1, 2, \ldots)$ is an open ν-cell, the boundary $B_\kappa = G_\kappa \cap \dot{W}_\kappa$ of W_κ in G_κ is according to the Phragmen-Brouwer theorem connected. Clearly $\overline{B}_\kappa$ lies in the compact set $\overline{G}_\kappa$ and hence is a continuum. Furthermore, $\overline{B}_\kappa \subset B \subset Q$. Now $\overline{B}_\kappa$ must be a planar set, else Q by containing a non-planar continuum would be nearly convex and its near complement null. Therefore B_κ is a planar portion cross-

cutting the convex set G_κ. We note that $B_\kappa \subset B_{\kappa+1}$. For any point $p \subset B_\kappa$ lies in $G_\kappa \cap \overline{W}_\kappa$ and hence in $G_{\kappa+1} \cap \overline{W}_{\kappa+1}$. If p were not in $B_{\kappa+1}$ it would lie in $W_{\kappa+1}$ and hence not in B. But this is impossible since $B_\kappa \subset B$. The union $B_\omega \subset B$ of the sets B_κ is then a planar portion crosscutting the union G of the sets G_κ. Consequently P is separated by the planar portion B_ω and hence is not semiconnected. This contradiction proves P connected.

This theorem suggests the question: Is a semiconnected quasiconvex set connected? The answer is no, as shown by the following example.

By using a procedure similar to that of Jones [17] a midpoint convex set Q dense in the Cartesian plane can be constructed having the property that Q intersects every perfect set not lying in a countable union of horizontal and vertical lines and having the further property that every horizontal line and every vertical line intersects Q in precisely one point. Clearly Q is semiconnected. However, the horizontal and vertical lines through any point not in Q form four complementary open quadrants one of which evidently contains no point of Q on its boundary. Thus Q can be separated by a right angle and hence is not connected; in fact, according to 4.6, Q has topological dimension zero.

From the theorem that a semiconnected near complement of a quasiconvex set is connected, we deduce the following three results.

THEOREM 5.6. *The near complement of a bounded semiconnected quasiconvex set is connected.*

Proof. The near complement P of a bounded semiconnected quasiconvex set Q is semiconnected and hence connected. For if, to the contrary, P is not semiconnected, it is a non-null set separated by a planar portion lying in Q. But then Q, being semiconnected, is, according to Theorem 5.3, nearly convex, whence P is null—a contradiction.

A set whose convex hull is the entire space X will be called totally unbounded.

THEOREM 5.7. *The near complement of a totally unbounded semiconnected quasiconvex set is connected.*

Proof. The near complement $P = CQ$ of a totally unbounded convex set Q is semiconnected and hence connected. For if, to the contrary, P is not semiconnected, it is a non-null set separated by some plane T, which, since Q is unbounded, lies in Q. Let f be a Δ reflection through a point of the non-null set P. Thus $f(T)$ is a plane lying in P and hence separating Q. This, however, is a contradiction, for Q is semiconnected.

We have shown that the near complement of a bounded or of a totally unbounded semiconnected quasiconvex set is connected. However, if the set is neither bounded nor totally unbounded its near complement may not be connected. For let Q be the intersection with the upper half plane $y > 0$ of the midpoint convex set consisting of all points (x, y) in the Cartesian plane such that $y > \varphi(x)$ where φ is a discontinuous solution of the functional equation

$\varphi(x+y)=\varphi(x)+\varphi(y)$. Now the set A of real numbers x for which $\varphi(x)\leqslant 0$ is everywhere dense; so for every $x \subset A$ the vertical line $y>0$ with abscissa x lies in Q. Therefore Q is a semiconnected midpoint convex set whose near interior, the upper half plane $y>0$, possesses only linear components. We see from 5.3 that Q is connected.

THEOREM 5.8. *If the near complement of a totally unbounded quasiconvex set contains a non-planar connected subset, it is connected.*

Proof. If the near complement $P = CQ$ of a totally unbounded Δ convex set Q contains a non-planar connected subset E, then P is semiconnected and hence connected. For if, to the contrary, P is not semiconnected, it is a non-null set separated by some plane T, which since Q is totally unbounded, lies in Q. Let G be the near interior of the non-planar set E. Then G is a non-null open set. Evidently a Δ^* contraction f toward a point of Q can be found such that the plane $f(T)$ intersects G. Thus $f(T)$ lies in Q and hence separates E. This, however, is impossible, for E is connected.

The following example shows that the near complement of a bounded quasiconvex set may possess exactly two non-planar components.

Let a and b be rationally independent real numbers. Then any rational linear combination x of a and b is uniquely expressible in the form $x = \alpha + \beta$ where α represents a rational multiple of a and β a rational multiple of b. Let Q be the set of all points (x, y) in the Cartesian plane such that x is expressible in the form $x = \alpha + \beta$ with $|\alpha| < 1$, $|\beta| < 1$, $|x| < 1$, and such that $-1 + |\beta| < y < 1 - |\alpha|$. It is easily verified that Q is a midpoint convex set whose near interior G is the square $|x| < 1$, $|y| < 1$. Moreover, the y-axis separates the near complement $G - Q$ of Q, but $G - Q$ is otherwise semiconnected. Thus $G - Q$ is not connected, but that part of it to either side of the y-axis is a nonlinear semiconnected and hence connected set.

History. A function φ defined for all real numbers satisfying the functional equation

$$\varphi(x+y) = \varphi(x) + \varphi(y) \tag{1}$$

will be called additive. In 1821 Cauchy [9] showed that any additive function φ is also rationally homogeneous, that is,

$$\varphi(\xi x) = \xi\varphi(x) \tag{2}$$

for all rational numbers ξ; whence he deduced that a continuous additive function φ is real homogeneous, that is, satisfies (2) for all real ξ and hence is of the form

$$\varphi(x) = x\varphi(1). \tag{3}$$

From (1) and (2) it follows that for any additive function φ we have

$$\varphi(\textstyle\sum \xi_\kappa x_\kappa) = \sum \xi_\kappa \varphi(x_\kappa), \tag{4}$$

the sum being finite and the ξ_κ being rational. In 1905 Hamel [14], using the then newly discovered well-ordering theorem of Zermelo, constructed a set H, now called a Hamel basis, with the property that every real number x can be represented uniquely (with the exception of zero coefficients) in the form

$$x = \sum \xi_\kappa x_\kappa, \tag{5}$$

where the sum is finite, the ξ_κ rational, and the x_κ belong to H. Thus we see from (4) that an additive function φ is exactly determined by its values on a Hamel basis H. If the functional values of φ are arbitrarily selected for x in H and determined for the remaining real numbers x by (4), then the resulting function φ is additive. It is continuous if and only if the ratio $\varphi(x)/x$ is constant as x ranges over the basis H. In this way Hamel completely solved the problem of the existence of discontinuous additive functions. Other interesting properties of Hamel bases and their application to discontinuous additive functions have been studied by Burstin [8], Sierpinski [30], and Jones [17, 18].

A function φ defined on an open interval of real numbers satisfying the functional inequality

$$\varphi(\tfrac{1}{2}\, x + \tfrac{1}{2}\, y) \leqslant \tfrac{1}{2}\, \varphi(x) + \tfrac{1}{2}\, \varphi(y) \tag{6}$$

will be called midpoint convex. Such functions were introduced in 1905 by Jensen [15, 16] who showed that any midpoint convex function φ is rationally convex, that is,

$$\varphi(\alpha x + \beta y) \leqslant \alpha\varphi(x) + \beta\varphi(y) \tag{7}$$

for all rational complementary ratios α and β; whence it follows that a continuous midpoint convex function is convex, that is, satisfies (7) for all real complementary ratios α and β.

Now it is easily seen from (4) that an additive function φ satisfies (7) with equality holding and hence is midpoint convex. Thus from the point of view of attaining generality it is desirable to consider the midpoint convex functions rather than the additive functions. Historically, however, results were first discovered for additive functions and then later extended to midpoint convex functions.

Generally speaking the problem was this: Find constraints, in themselves very weak, which, when placed on an additive or midpoint convex function, are sufficiently strong to force that function to be continuous; that is: What pathological properties do the discontinuous additive and midpoint convex functions possess?

Some density properties of additive functions were noted in 1875 by Darboux [10, 11] who showed that an additive function is continuous if it is bounded above or below on some interval. In his paper on the generation of discontinuous additive functions Hamel [14] pointed out that the graph of such a function is everywhere dense in the plane. In 1915, Bernstein and Doetsch [6] showed that the graph of a discontinuous midpoint convex function is dense above some convex function ($-\infty$ being allowed).

Measure properties of additive and midpoint convex functions have been extensively investigated. The first result in this direction, namely, that a measurable additive function is continuous, was discovered in 1913 by Fréchet [12]. This same theorem has since been proved many times: in 1920 by Sierpinski [31] and by Banach [2], in 1936 by Kac [19], in 1945 by Alexiewicz and Orlicz [1], and in 1947 by Kestleman [20]. It was somewhat generalized in 1924 by Sierpinski [34], who observed that an additive function majorized by a measurable function is continuous. That a measurable midpoint convex function is continuous was shown by Blumberg [7] in 1919 and in 1920 by Sierpinksi [32]. It should be mentioned that these measure results are closely connected with the work of Steinhaus [35] on the distances between points of a set.

These researches into measure and density properties culminated in 1929 in two papers by Ostrowski [22, 23] which include practically all the previous results. In one paper [22] Ostrowski showed that a midpoint convex function bounded above on a set of positive measure is continuous; and in the other paper [23] that the x-projection of the set of those graph points of a discontinuous midpoint convex function which lie in any plane neighbourhood above the lower bounding curve of the function has positive outer measure and zero inner measure.

The connectivity properties of graphs of discontinuous additive functions were first studied in 1942 by Jones [17, 18], who showed that every such graph is either connected or totally disconnected, and that it is connected if and only if it intersects every non-vertical continuum. Jones also pointed out how many pathological properties that connected sets may possess can be exhibited by the graphs of discontinuous additive functions or by sets closely related with such graphs, thus unifying and simplifying a large collection of examples scattered throughout the literature.

Other papers not mentioned in this historical survey which appear in our bibliography are: [4, 5, 13, 21, 24, 25, 26, 27, 28, 33]. An excellent account of the development of convex functions and sets (including midpoint convexity) and their generalizations may be found in a recent article by Beckenbach [3].

Conclusion. Our point of view throughout this paper has been on sets rather than on functions. Theorems concerning quasiconvex sets are applicable to the study of functions; for the graph of an additive function is, as we have already mentioned, midpoint convex, and the set of points (x, y) such that $y \geqslant \varphi(x)$, where φ is a midpoint convex function, is also a midpoint convex set. With few exceptions the theorems concerning additive and midpoint convex functions can be deduced from results on quasiconvex sets; though it is not generally conversely true that the set results can be made to follow from the function theorems.

References

[1] A. Alexiewicz and W. Orlicz, *Remarque sur l'équation fonctionelle,* $f(x + y) = f(x) + f(y)$. Fund. Math., vol. 33 (1945), 314-315.

[2] S. Banach, *Sur l'équation fonctionelle* $f(x+y) = f(x) + f(y)$, Fund. Math., vol. 1 (1920), 123-124.
[3] E. F. Beckenbach, *Convex Functions*, Bull. Amer. Math. Soc., vol. 54 (1948), 439-460.
[4] E. F. Beckenbach and R. H. Bing, *On generalized convex functions*, Trans. Amer. Math. Soc., vol. 58 (1945), 220-230.
[5] F. Bernstein, *Über das Gausssche Fehlergesetz*, Math. Ann., vol. 64 (1907), 417-448.
[6] F. Bernstien and G. Doetsch, *Zur Theorie der konvexen Funktionen*, Math. Ann., vol. 76 (1915), 514-526.
[7] H. Blumberg, *Convex functions*, Trans. Amer. Math. Soc., vol. 20 (1919), 40-44.
[8] C. Burstin, *Die Spaltung des Kontinuum in c in L Sinne nichtmessbare Mengen*, Sitzber. Akad. Wiss. Vienna, Math-nat. Kl., Abt, IIa, vol. 125 (1916), 209-317.
[9] A. L. Cauchy, *Cours d'analyse de l'École Royale Polytechnique*, part 1, Analyse algebrique (Paris, 1821).
[10] G. Darboux, *Sur la composition des forces en statique*, Bull. Sci. Math., vol. 9 (1875), 281-288.
[11] ——— *Sur la théorème fondamental de la géometrie projective*, Math. Ann., vol. 17 (1880), 55-61.
[12] M. Fréchet, *Pri la funkcia ekvacio* $f(x+y)=f(x)+f(y)$, Ens. Math., vol. 15 (1913), 390-393.
[13] G. Hamel, *Über die Zusammensetzung von Vektoren*, Zeit. Math. Phys., vol. 49, 362-371.
[14] ——— *Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung:* $f(x)+f(y)=f(x+y)$, Math. Ann., vol. 60 (1905), 459-462.
[15] J. L. W. V. Jensen, *Om konvexe Funktioner og Uligheder mellem Middelvaerdier*, Nyt. Tidsskrift for Mathematik, vol. 16B (1905), 49-69.
[16] ——— *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Math., vol. 30 (1906), 175-193.
[17] F. B. Jones, *Connected and disconnected plane sets and the functional equation* $f(x)+f(y)=f(x+y)$, Bull. Am. Math. Soc., vol. 48 (1942), 115-120.
[18] ——— *Measure and other properties of a Hamel basis*, Bull. Amer. Math. Soc., vol. 48 (1942), 472-481.
[19] M. Kac, *Une remarque sur les équations fonctionelles*, Comm. Math. Helv., vol. 9 (1936/37), 170-171.
[20] H. Kestelman, *On the functional equation* $f(x+y)=f(x)+f(y)$, Fund. Math., vol. 34 (1947), 144-147.
[21] H. Lebesgue, *Sur les transformations ponctuelles, transformant les plans en plans, qu'on peut définir par des procédés analytiques*, Atti R. Ac. Sc. Torino, vol. 42 (1907).
[22] A. Ostrowski, *Über die Funktionalgleichung der Exponentialfunktion und verwandte Funktionalgleichung*, Jber. Deut. Math. Ver., vol. 38 (1929), 54-62.
[23] ——— *Zur Theorie der konvexen Funktionen*, Comm. Math. Helv., vol. 1 (1929), 157-159.
[24] V. Ramaswami, *On the continuity of convex functions*, J. Benares Hindu Univ., vol. 7, part 1 (1943), 180-181.
[25] S. Ruziewicz, *Une application de l'équation fonctionelle* $f(x+y)=f(x)+f(y)$ *à la décomposition de la droite en ensembles superposables, non mesurables*, Fund. Math., vol. 5 (1924), 92-95.
[26] ——— *Contribution à l'étude des ensembles des distances de points*, Fund. Math., vol. 7 (1925), 141-143.
[27] R. Schimmack, *Über die axiomatische Begründung der Vektoraddition*, Gött. Nach. (1903), 317-325.
[28] ——— *Axiomatische Untersuchungen über die Vektoraddition*, Nova Acta Ac. Leop., vol. 90, 5-104.
[29] F. Schur, *Über die Zusammenhang von Vektoren*, Zeit. Math. Phys., vol. 49, 352-361.
[30] W. Sierpinski, *Sur la question de la mesurabilité de la base de M. Hamel*, Fund. Math., vol. 1 (1920), 105-111.

[31] ——— *Sur l'équation fonctionelle $f(x+y)=f(x)+f(y)$*, Fund. Math., vol. 1 (1920), 116-122.
[32] ——— *Sur les fonctions convexes mesurables*, Fund. Math., vol. 1 (1920), 125-129.
[33] ——— *Sur l'ensemble de distance entre les points d'un ensemble*, Fund. Math., vol. 7 (1925), 144-148.
[34] ——— *Sur une propriété des fonctions de M. Hamel*, Fund. Math., vol. 5 (1924), 334-336.
[35] H. Steinhaus, *Sur les distances des points des ensembles de mesure positive*, Fund. Math. vol. 1 (1920), 93-104.

U.C.L.A. and
Indiana University

ON THE VECTOR SUM OF CONTINUA

J. W. GREEN AND W. GUSTIN

IN this note we investigate certain properties of a set formed as the vector sum of continua. Our interest in this subject arose in connection with the preceding paper *Quasiconvex sets* where we use, but do not prove, item 3 below.

Let X be a normed real vector space of dimension ν with the ν vectors a_λ as a basis (λ representing a variable index ranging over the ν indices $1, \ldots, \nu$). The parallelepipedal lattice consisting of all integral linear combinations $a = \sum \alpha_\lambda a_\lambda$ of the basis vectors a_λ, the coefficients α_λ being integers, will be denoted by A.

Consider ν continua Q_λ in X such that the λth continuum Q_λ contains the origin 0 and the basis vector a_λ. Let $Q = \sum Q_\lambda$ be the vector sum of these ν continua Q_λ, that is, the set of all vector sums $q = \sum q_\lambda$ with $q_\lambda \subset Q_\lambda$. A simple example of such a set is the solid parallepiped $P = \sum P_\lambda$ where P_λ is the line segment joining 0 and a_λ.

We shall prove that any vector sum Q, formed as above described, possesses the following properties:

(1) Q is a continuum, (2) $X = A + Q$,

(3) Q has interior points, (4) $\mu(P) \leqslant \mu(Q)$,

where μ is a measure on the space X invariant under translation.

1. We are to show that Q is a continuum: compact and connected.

We first demonstrate that Q is compact. To this end let $q^\gamma = \sum q_\lambda{}^\gamma$ with $q_\lambda{}^\gamma \subset Q_\lambda$ be a sequence of points in Q, γ running through the sequence Γ of positive integers. It is required to find a point $q \subset Q$ with $q^\gamma \to q$ as γ runs through some subsequence of Γ. Consider the sequence of points $q_1{}^\gamma$ in Q_1 as γ runs through Γ. Since Q_1 is compact a point $q_1 \subset Q_1$ and a subsequence Γ_1 of Γ exists with $q_1{}^\gamma \to q_1$ as γ runs through Γ_1. Since Q_2 is compact a point $q_2 \subset Q_2$ and a subsequence Γ_2 of Γ_1 exists with $q_2{}^\gamma \to q_2$ as γ runs through Γ_2. Continuing this process recursively to the νth stage we obtain ν points $q_\lambda \subset Q_\lambda$ and ν sequences $\Gamma_1, \ldots, \Gamma_\nu$ each a subsequence of the preceding one such that $q_\lambda{}^\gamma \to q_\lambda$ as γ runs through Γ_λ and hence also as γ runs through Γ_ν. Putting $q = \sum q_\lambda$ we have $q \subset Q$, since $q_\lambda \subset Q_\lambda$, and

$$q^\gamma = \sum q_\lambda{}^\gamma \to \sum q_\lambda = q$$

as γ runs through Γ_ν. This proves Q compact.

We next demonstrate that Q is connected by constructing for each point

Received July 1, 1949.

$q \subset Q$ a connected set C containing the origin 0 and the point q. Let $q = \sum q_\lambda$ with $q_\lambda \subset Q_\lambda$; and define the points c_λ and sets C_λ as follows:

$$c_0 = 0, \qquad c_\lambda = c_{\lambda-1} + q_\lambda, \qquad C_\lambda = c_{\lambda-1} + Q_\lambda, \qquad C = \cup C_\lambda.$$

We note that C_λ, being a translation of Q_λ, is connected and, since Q_λ contains 0 and q_λ, contains $c_{\lambda-1}$ and c_λ. Consequently the set C is connected and contains $c_0 = 0$ and $c_\nu = q$. This proves Q connected.

2. Since the vectors a_λ constitute a basis for the ν dimensional vector space X, every vector x in the space may be expressed uniquely in the form

$$x = \sum \xi_\lambda(x) a_\lambda,$$

the coordinate functionals $\xi_\lambda(x)$ being linear. Thus the function

$$\omega(x) = \max |\xi_\lambda(x)|$$

has the properties of a norm. Now any two norm topologies on X are topologically equivalent, so we shall for convenience assume that ω is the norm on X and write $\omega(x) = |x|$. Observe that with this norm we have $|p| \leqslant 1$ for all $p \subset P$.

Every real number ξ can be uniquely partitioned into an integer α and a remainder θ with $0 \leqslant \theta < 1$ so that $\xi = \alpha + \theta$. Let this partition for the coordinate functional $\xi_\lambda(x)$ be

$$\xi_\lambda(x) = \alpha_\lambda(x) + \theta_\lambda(x)$$

and define

$$a(x) = \sum \alpha_\lambda(x) a_\lambda, \qquad p(x) = \sum \theta_\lambda(x) a_\lambda .$$

Then $a(x) \subset A$, $p(x) \subset P$, and

$$x = a(x) + p(x),$$

(which shows $X = A + P$).

Fix the positive number $\epsilon > 0$. Since Q_λ is connected, any two of its points may be connected in it by an ϵ chain. Thus there exists an ϵ chain C_λ^ϵ of points of Q_λ running from 0 to a_λ. Consider the path Q_λ^ϵ obtained by drawing the line segments joining consecutive points along the chain C_λ^ϵ. This path begins at 0 and ends at a_λ, and is at all of its points within ϵ distance of some point of Q_λ, namely a point of C_λ^ϵ. Clearly a continuous mapping q_λ^ϵ can be constructed which maps the closed real unit interval $I: 0 \leqslant \theta \leqslant 1$ onto the path Q_λ^ϵ so that $q_\lambda^\epsilon(0) = 0$ and $q_\lambda^\epsilon(1) = a_\lambda$. Let $\xi = \alpha + \theta$ be the partition of the real number ξ into its integral part α and remainder θ; and define

$$f_\lambda^\epsilon(\xi) = f_\lambda^\epsilon(\alpha + \theta) = \alpha a_\lambda + q_\lambda^\epsilon(\theta).$$

We note that the mapping f_λ^ϵ is continuous for all real ξ. This is obvious for non-integral ξ and also for integral ξ approached from above, since q_λ^ϵ is continuous on I and the integral part of ξ becomes constant. Furthermore for

integral $\xi = a + 1$ (not a partition) f_λ^ϵ is also continuous from below; for, as $\theta \to 1$ from below, we have

$$f_\lambda^\epsilon(a + \theta) = a a_\lambda + q_\lambda^\epsilon(\theta) \to a a_\lambda + q_\lambda^\epsilon(1) = (a + 1)a_\lambda = f_\lambda^\epsilon(a + 1).$$

Thus f_λ^ϵ is continuous.

Now define the mapping f^ϵ of the space X into itself as follows:

$$f^\epsilon(x) = \sum f_\lambda^\epsilon(\xi_\lambda(x)) = a(x) + q^\epsilon(x)$$

where

$$q^\epsilon(x) = \sum q_\lambda^\epsilon(\theta_\lambda(x)).$$

Since f_λ^ϵ and ξ_λ are continuous mappings, f^ϵ is also continuous. Every point of Q_λ^ϵ is within ϵ distance of some point of Q_λ, so we see that $q^\epsilon(x)$ is within $\nu\epsilon$ distance of some point of Q. Now

$$x - f^\epsilon(x) = p(x) - q^\epsilon(x),$$

so

$$|x - f^\epsilon(x)| \leqslant 1 + \nu\epsilon + \delta$$

where $\delta = \max |q|$ $(q \subset Q)$. Thus $x - f^\epsilon(x)$ is uniformly bounded for all x in X.

Let S be the open unit sphere: $|s| < 1$. The mapping h defined by the formula

$$s = h(x) = \frac{x}{1 + |x|}$$

is a homeomorphism contracting X onto S whose inverse mapping is

$$x = h^{-1}(s) = \frac{s}{1 - |s|}, \quad |s| < 1.$$

With several applications of the triangle inequality it may be shown that the homeomorphism h satisfies the following norm inequality which we shall call the h-inequality:

$$|h(x) - h(y)| \leqslant |x - y| \cdot (1 - |h(x)| \cdot |h(y)|).$$

Consider now the mapping g^ϵ of the closed sphere $\overline{S} : |s| \leqslant 1$ into itself defined by

$$g^\epsilon(s) = \begin{cases} hf^\epsilon h^{-1}(s), & |s| < 1, \\ s, & |s| = 1. \end{cases}$$

This mapping g^ϵ is, as we shall demonstrate, continuous on $\overline{S}$. It is clearly continuous when confined either to S or to its boundary. Thus it remains to show that $g^\epsilon(s^\gamma) \to g^\epsilon(s)$ as $s^\gamma \to s$, the points s^γ being in S and the point s on the boundary of S. Let

$$x^\gamma = h^{-1}(s^\gamma),$$

$$y^\gamma = f^\epsilon(x^\gamma),$$

$$d^\gamma = s^\gamma - g^\epsilon(s^\gamma) = h(x^\gamma) - h(y^\gamma).$$

Since, as we have already noted, the vectors $x^\gamma - f^\epsilon(x^\gamma) = x^\gamma - y^\gamma$ are uniformly bounded independently of γ, and $s^\gamma \to s$ (so that $|s^\gamma| \to |s| = 1$), we have $|x^\gamma| \to \infty$ and $|y^\gamma| \to \infty$. Therefore $|h(x^\gamma)| \to 1$ and $|h(y^\gamma)| \to 1$, so it follows from the h-inequality that $|d^\gamma| \to 0$ and hence that $d^\gamma \to 0$. Consequently

$$g^\epsilon(s^\gamma) - g^\epsilon(s) = s^\gamma - s - d^\gamma \to 0,$$

which proves g^ϵ to be a continuous mapping of the closed sphere $\bar{S}$ into itself leaving the boundary fixed. According to a variant form of the Brouwer fixed point theorem for the closed sphere $\bar{S}$ the mapping g^ϵ is then an onto mapping: $g^\epsilon(\bar{S}) = \bar{S}$, whence $g^\epsilon(S) = S$. Therefore f^ϵ is also an onto mapping:

$$f^\epsilon(X) = f^\epsilon h^{-1}(S) = h^{-1}g^\epsilon(S) = h^{-1}(S) = X.$$

Our present task is to show that $X = A + Q$; that is, for any point $x \subset X$ the equation $x = a + q$ is solvable with $a \subset A$ and $q \subset Q$. Choose a sequence of positive numbers ϵ^γ such that $\epsilon^\gamma \to 0$ as γ runs through the sequence Γ of positive integers. Since $f^{\epsilon^\gamma}(X) = X$ there exists a point x^γ such that

$$x = f^{\epsilon^\gamma}(x^\gamma) = a(x^\gamma) + q^{\epsilon^\gamma}(x^\gamma).$$

Put $a(x^\gamma) = a^\gamma$; and, since $q^{\epsilon^\gamma}(x^\gamma)$ is within $\nu\epsilon^\gamma$ distance of Q, replace it by $q^\gamma + e^\gamma$, where $q^\gamma \subset Q$ and $|e^\gamma| < \nu\epsilon^\gamma$. Thus

$$x = a^\gamma + q^\gamma + e^\gamma.$$

Since the q^γ and e^γ are bounded so also are the a^γ. But any bounded subset of A is finite so a^γ is constant, say a, for infinitely many integers γ, say for the sequence Γ_a. Q being compact, a point $q \subset Q$ and a subsequence Γ_{aq} of Γ_a exist with $q^\gamma \to q$ as γ runs through Γ_{aq}. Therefore

$$x = a^\gamma + q^\gamma + e^\gamma \to a + q$$

as γ runs through Γ_{aq}, since $a^\gamma = a$, $q^\gamma \to q$, and $e^\gamma \to 0$. This proves $x = a + q$.

3. We have shown that $X = A + Q$. Thus the space X is the union as a ranges over the countable set A of the translates $a + Q$ of Q. Since the entire space X is of second category, at least one of these translates of Q, and hence Q itself, is somewhere dense. Therefore the set Q, being closed, contains interior points.

4. The set Q is closed and hence measurable, so every translate $a + Q$ of Q is measurable, and has the measure $\mu(Q)$; similarly every translate $a + P$ of P is measurable and has measure $\mu(P)$.

Define A^β for each integer $\beta \geqslant 0$ to be that subset of A consisting of the $(2\beta + 1)^\nu$ integral linear combinations $a = \sum \alpha_\lambda a_\lambda$ of the basis vectors a_λ, the coefficients α_λ being integers such that $|\alpha_\lambda| \leqslant \beta$. Observe that by our selection of norm we have $|a| \leqslant \beta$ for every $a \subset A^\beta$. Let $Q^\beta = A^\beta + Q$ and $P^\beta = A^\beta + P$.

The set Q^β is the union as a ranges over the set A^β of the $(2\beta + 1)^\nu$ measurable sets $a + Q$ each having measure $\mu(Q)$. Therefore Q^β is measurable and

$$\mu(Q^\beta) \leqslant (2\beta + 1)^\nu \mu(Q).$$

Every plane, being closed, is measurable. Suppose that every plane has measure zero. The intersection of any two translates of P by distinct vectors of A being a planar set (possibly null) then has measure zero. Two such sets may be called μ-disjoint. Consequently P^β is the union as a ranges over A^β of the $(2\beta + 1)^\nu$ measurable pairwise μ-disjoint sets $a + P$ each having measure $\mu(P)$; so P is measurable and has measure

$$\mu(P^\beta) = (2\beta + 1)^\nu \mu(P).$$

We now show that $P^\beta \subset Q^{\beta+\gamma}$, where γ is any fixed integer $\geqslant \delta + 1$ and $\delta = \max |q|$ $(q \subset Q)$. Suppose, to the contrary, that some point x exists with $x \subset P^\beta$ and $x \not\subset Q^{\beta+\gamma}$. Since $x \subset P^\beta$ we have $x = a_p + p$ where $a_p \subset A^\beta$ and $p \subset P$, whence

$$|x| = |a_p + p| \leqslant |a_p| + |p| \leqslant \beta + 1.$$

Now $X = A + Q$ so $x = a_q + q$ where $a_q \subset A$ and $q \subset Q$. However $a_q \not\subset A^{\beta+\gamma}$ since $x \not\subset Q^{\beta+\gamma}$, so

$$|x| = |a_q + q| \geqslant |a_q| - |q| > \beta + \gamma - \delta \geqslant \beta + 1.$$

in contradiction to the preceding inequality. This contradiction proves $P^\beta \subset Q^{\beta+\gamma}$. Therefore

$$(2\beta + 1)^\nu \mu(P) = \mu(P^\beta) \leqslant \mu(Q^{\beta+\gamma}) \leqslant (2\beta + 2\gamma + 1)^\nu \mu(Q).$$

Dividing this inequality through by $(2\beta + 1)^\nu$ and letting $\beta \to \infty$ we obtain the desired inequality $\mu(P) \leqslant \mu(Q)$.

If, finally, some plane has positive measure, then it is possible by suitable translation to insert into any sphere infinitely many disjoint parallel planar portions all having the same positive measure, so that every set with interior points, in particular P and Q, has infinite measure.

U.C.L.A. and
Indiana University